BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

FILE

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N A2

(11) International Publication Number: **WO 96/34088**

(43) International Publication Date: 31 October 1996 (31.10.96)

(21) International Application Number: PCT/US96/03466

(22) International Filing Date: 15 March 1996 (15.03.96)

(30) Priority Data:
08/406,186 16 March 1995 (16.03.95) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/406,186 (CIP)
Filed on 16 March 1995 (16.03.95)

(71) Applicant *(for all designated States except US)*: COLD SPRING HARBOR LABORATORY [US/US]; P.O. Box 100, Cold Spring Harbor, NY 11724 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: COLASANTI, Joseph, J. [CA/US]; 19 West 10th Street, Huntington Station, NY 11746 (US). SUNDARESAN, Venkatesan [IN/US]; 69 Inwood Road, Port Washington, NY 11050 (US).

(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).

(81) Designated States: CA, JP, MX, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: CONTROL OF FLORAL INDUCTION IN PLANTS AND USES THEREFOR

(57) Abstract

The *Id* gene which controls flower evocation in maize plants is described. The maize nucleic acid is similar to that of genes encoding zinc-finger regulatory proteins in animals. Methods of isolation or preparation of other regulatory protein genes in plants and their uses are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia
AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

Control of Floral Induction in Plants and Uses Therefor

Background of the Invention

Higher plants have a life cycle that consists of a period of vegetative growth followed by reproductive development. Reprcduction in angiosperms is a developmental process that begins with floral induction (evocation). This is the point in time at which the shoot apical meristem, the set of dividing cells that gives rise to most of the plant parts above the roots, stops making leaves and starts making flowers. Bernier, G. (1988) The control of floral evocation and morphogenesis. *Ann. Rev. Plant. Physiol. Plant Molec. Biol.* 39:175-219. Almost nothing is known, however, about the molecular and genetic controls that induce a plant to flower.

There is a great need for more information about the regulatory elements in plants. Increased knowledge of these elements would significantly improve our understanding of the underlying mechanism by which genes induce reproductive development in plants.

Summary of the Invention

This invention identifies and provides isolated DNA which comprises an *Id* gene of a maize plant, or a portion thereof, which demonstrates *Id* gene function. The invention further provides RNA encoded by the DNA of the *Id* or *id** alleles and portions thereof, and antisense (complementary) DNA and/or RNA or portions thereof. Nucleic acids, referred to as *Id* homologues or equivalents, which a) show greater than 50% homology or that hybridize under moderate stringency conditions to the zinc finger region of the *Id* gene or b) show a 70% or greater homology or that hybridize under moderate stringency conditions to the *Id* gene and demonstrate *Id*-type (initiation of

reproduction phase) function are also encompassed by this invention. Nucleic acid probes and primers to detect and/or amplify regulatory genes in other plants are included as well. Thus, the DNA of this invention comprises an *Id* gene, or a portion thereof, the *Id* gene comprising all or a portion of SEQ ID NO:1, or homologous DNA.

The present invention further encompasses polypeptides which are *Id* proteins or portions of an *Id* protein of plant origin, including the polypeptides herein described. *Id* proteins from all plant species or homologues demonstrating a similar regulatory function (reproductive induction) are encompassed by this invention and the term *Id* protein as used herein. Amino acid sequences that demonstrate 80% or greater homology to the amino acid sequences described herein are considered homologous polypeptides.

In another aspect, this invention relates to antibodies which bind the polypeptides described herein. Such antibodies can be used to locate sites of regulatory activity in plants. Fusion proteins comprising the *Id* protein and an additional peptide, such as a protein tag, can also be used to detect sites of *Id* protein/protein interaction in plants.

In a further aspect, this invention provides methods for producing plants with selected times of transition from the vegetative to the flowering stage. Applicants have created a new allele of the *id* gene, *id**, which, when an active *Ac* transposable element is present, causes plants to stop vegetative growth and to flower earlier than do other *id* mutants. As shown herein, the *id*/id** plants with an active *Ac* element exhibit fewer vegetative nodes and flower earlier than *id*/id** plants without an *Ac* element or plants encoding the *id* allele.

The present invention relates to a new mutant of the *id* gene which encodes a product that alters flower

induction in plants and provides a nucleotide sequence of part of the *Id* *Sac*I 4.2 kb fragment derived from maize Chromosome 1. Also included is DNA which hybridizes under high stringency conditions to the *Sac*I fragment or a portion thereof and an RNA transcribed from or corresponding to either of said aforementioned DNA. Preferably the DNA is that shown in Figure 4 (SEQ ID NO:3).

In another aspect, this invention provides methods for producing new *id* alleles and methods for detecting other *Id* alleles or other regulatory genes in plants. Homologues of the *Id* gene can be identified throughout the plant kingdom, including the multicellular and unicellular algae.

In yet another aspect of this invention are provided plants, seeds, plant tissue culture, and plant parts which contain DNA comprising an altered or exogenously introduced *Id* allele or portion of an *Id* allele that alters the timing of flower induction in the subsequent growth of the plant, seeds, plant tissue culture, and/or plant part.

The present invention also relates to transgenic plants in which the time of floral evocation is altered. Transgenic plants are provided in which the time period from germination to flowering is shorter than it is in the corresponding naturally-occurring or wild type (native) plant. Alternatively, plants are provided in which flowering is delayed or absent. As used herein, the term transgenic plants includes plants that contain either DNA or RNA which does not naturally occur in the wild type (native) plant or known variants, or additional or inverted copies of the naturally-occurring DNA and which is introduced as described herein, and any of the above-described alterations which result in plants having altered floral evocation times. Such transgenic plants include, in one embodiment, transgenic plants which are angiosperms, both monocotyledons and dicotyledons. Transgenic plants include those into which DNA has been introduced and their

progeny, produced from seed, vegetative propagation, cell, tissue or protoplast culture, or the like.

Transgenic plants of the present invention contain DNA which encodes all or a portion of a protein essential for floral evocation and, when present in plant cells, results in altered floral evocation, either earlier cessation of vegetative growth and initiation of flowering than in untransformed plants of the same variety, or in later flowering or the absence of floral induction. The DNA can be exogenous DNA in a sense or antisense orientation which encodes a protein required for floral induction or exogenous DNA which has been altered in such a manner that it encodes an altered form of a protein required for floral induction. Directed or targeted mutagenesis of a plant's endogenous DNA responsible for initiation of flowering can also result in altered floral induction. Exogenous DNA encoding an altered protein required for floral evocation and endogenous DNA required for floral evocation which has been mutated by directed mutagenesis differ from the corresponding wild type (naturally-occurring) DNA in that these sequences contain a substitution, deletion or addition of at least one nucleotide and encode proteins which differ from the corresponding wild type protein by at least one amino acid residue. (As used herein, the term "nucleotide" is used interchangeably with "nucleic acid".) Insertion of genetic elements, such as *Ds* sequences with or without active *Ac* sequences, are of particular use.

Exogenous DNA is introduced into plant cells of the target plant by well-known methods, such as Agrobacterium-mediated transformation, microprojectile bombardment, microinjection or electroporation (see-below). Such cells carrying the introduced exogenous DNA or endogenous *Id* DNA mutated by direct mutagenesis can be used to regenerate transgenic plants which have altered floral induction, therefore becoming sources of additional plants either

through seed production or non-seed asexual reproductive means (i.e., cuttings, tissue culture, and the like).

The present invention also relates to methods of producing plants with altered floral induction times, exogenous DNA or RNA whose presence in a plant results in altered floral induction, and vectors or constructs which include DNA or RNA useful for producing recombinant plants with altered floral development. Seeds produced by plants which contain exogenous DNA or RNA encoding a protein which is required for floral induction, such as *Id* DNA in the sense orientation or exogenous DNA which has been altered in such a manner that it encodes an altered form of a protein required for floral development, such as altered *id** DNA, are also the subject of the present invention.

The work described herein makes available an *Id* gene, the genomic sequence, or a portion thereof, which has been determined by the Applicants, and which has an important role in the induction of flowering of plants. The gene is derived from a monocot, specifically, maize, one of the most commercially valuable grasses. The polypeptide encoded by this gene is a regulatory protein that causes a switch from vegetative growth to the development of reproductive organs in maize. In addition, in maize as in many other plants, the effects of this protein marks the beginning of senescence in these plants.

Corn requires more rainfall than wheat and most maize cultivars need a long growing season. The work described herein also makes it possible to grow maize and other latitude-dependent plants which require long growing seasons before flowering can take place to be grown in geographic regions with short growing seasons. Thus, the plants can be induced to flower and set seed prior to the first frost. Similarly, flower induction can be prolonged for short-season plants grown in areas with long periods of warm weather. As a result of the extra vegetative mass and

carbohydrate, these plants can produce more and/or larger flowers and, consequently, more seed. Or, plants can even be prevented from flowering, thus providing nutritious silage biomass.

In another aspect, this invention provides a means to eliminate the need for detasseling in the production of maize and sorghum hybrids.

Brief Description of the Figures

Figure 1 is a map of Chromosome 1 showing the location of the *indeterminate* and *Bz2* (*bronze* kernel pigmentation) genes, and the site of transposon insertion for *Ds2*.

Figure 2 is the genomic sequence (SEQUENCE ID NO:1) comprising DNA of the *Id* gene.

Figures 3A-3F is the genomic sequence of Figure 2 and the deduced amino acid sequences a, b and c (SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:2, respectively). The *Ds2* transposon insertion occurs at nucleotide 168.

Figure 4 is a restriction map of the conserved motif of the 4.2 kb *Sac*I fragment which includes a portion of the *Id* gene. The location of the *Ds2* transposon insertion and the genomic sequence (SEQUENCE ID NO:3) between restriction sites *Nsi*I and *Sac*I are shown.

Figure 5 shows the polypeptide sequence (SEQUENCE ID NO:4) encoded by SEQUENCE ID NO:3. (The ORF showing part of the protein sequence is also the portion showing homology to Zn-finger proteins.)

Figure 6 is a comparison of the maize *Id* gene ORF to known zinc-finger proteins of eukaryotic animal species.

Figure 7 shows the frame shifts produced by the excision of *Ds2* from the *Id* gene ORF, resulting in the two null mutants, *id-X1* and *id-X2*.

Figures 8A-8B depict schematic representations of *Id* antisense constructs in which a weak promoter is fused with

the *Id* cDNA for production of transgenic (Figure 8A) monocots or (Figure 8B) dicots to delay flowering in an early flowering line.

Figure 9A-9B depict schematic representations of *Id* sense construct in which a constitutive promoter is fused with the *Id* cDNA for production of transgenic (Figure 9A) monocots or (Figure 9B) dicots to induce early flowering in a late flowering line.

Figure 10A-10B depicts schematic representations of *Id* antisense constructs in which a drought induced promoter is fused with the *Id* cDNA for production of transgenic (Figure 10A) monocots or (Figure 10B) dicots to delay flowering in response to drought.

Figure 11A-11D depicts schematic representations of *Id* antisense constructs in which a GAL4 binding site (GB) is fused with the *Id* cDNA in a monocot (11A) or a dicot (11B), and a GAL4 gene is fused with a strong (CaMV 35s) or weak promoter in a monocot (11C) or a dicot (11D), for production of transgenic plants in which flowering is absent or delayed.

Detailed Description of the Invention

During reproductive growth the plant enters a program of floral development that culminates in fertilization, followed by the production of seeds. Senescence may or may not follow. A maize plant (or its close relative, sorghum) is normally programmed to generate a particular number of vegetative structures (e.g. leaves), followed by reproductive structures (flowers), and to eventually undergo senescence of the plant. Maize (*Zea mays*) plants that are homozygous for the *indeterminate* (*id*) mutation of the *Id* gene, however, are defective in the execution of this program and exhibit several developmental phenotypes: 1) The vegetative to reproductive transition is altered such that the vegetative phase is prolonged, resulting in

plants with an extensive (or indeterminate) lifespan; i.e., they flower much later than normal plants, or not at all. 2) The vegetative phase expands into the reproductive phase of development and causes abnormal flower development; i.e., the female flower (ear) exhibits vegetative characteristics and is usually sterile, and the male flower (tassel) can undergo a complete developmental reversion such that new vegetative shoots emerge from tissues that have characteristics of floral tissue. In the latter case, terminally differentiated cells that comprise floral tissues redifferentiate into vegetative tissue and resume proliferative growth. Singleton, W.R., *J. Heredity*, 37:61-64 (1946); Galinat, W.C. and Naylor, A.W. (1951) *Am. J. Bot.* 38:38-47. These phenotypes suggest that the function of the normal *Id* gene is to suppress vegetative growth and signal the beginning of reproductive growth at a specific time during the life cycle of the plant. Loss of *Id* function results in the failure to make this transition and causes prolonged vegetative development.

Normal *Id* function, therefore, is important in the vegetative to reproductive transition in maize; i.e., floral induction or evocation. Genetic and molecular data suggest that the *Id* gene encodes a regulatory protein that plays a crucial role in the switch from vegetative to reproductive development in maize and other plants. Understanding the mechanism of this regulation provides a basis for producing specialized plants designed to flower and produce seed independent of native internal controls or environmental effects. In fact, it is possible that the same mechanism utilizing a homologue of the *Id* gene controls spore production in non-seed plants, such as the algae.

The term "*Id*" means the normal (wild-type) gene of maize; whereas, "*id*" refers to an altered (mutant) form of the *Id* gene. Isolated DNA of plant which encodes

polypeptides which trigger initiation of the reproductive phase in the plant can be genomic or cDNA. DNA included in the present invention is from monocots, grasses; specifically described is the *Id* gene from maize.

Applicants have created a new allele of the *id* mutation that results from the disruption of normal *Id* gene function by the insertion of the 1.3 kb transposable element *Dissociation (Ds)* into the gene. A clone containing a portion of the mutated *id* gene, *id**, was then isolated by the technique of transposon tagging using *Ds* as the tag. Hake, *et al.*, *EMBO J.*, 8:15-22 (1989); Federoff *et al.* (1984) *PNAS* 81:3825-3829. Preliminary sequence analysis of a portion of the gene (*id** and *Id*) indicates that *Id* contains regions that are homologous to a class of transcription factor found in all eukaryotic organisms.

A transposable genetic element (transposon) is a piece of DNA that moves from place to place in an organism's genome. It is excised from one site and inserted at another site, either on the same chromosome or on a different one. The movement of a transposable element can generate mutations or chromosomal rearrangements and thus affect the expression of other genes.

Transposons *Ac* and *Ds* constitute a family of related transposable elements present in maize. Fedoroff, N. (1989) *Maize Transposable Elements*. In Mobile DNA, M. Howe and D. Berg, eds, Washington: ASM press. *Ac* is able to promote its own transposition or that of *Ds* to another site, either on the same chromosome or on a different one. *Ds* cannot move unless *Ac* is present in the same cell. *Ac* is an autonomous transposable element and *Ds* is a nonautonomous element of the same family.

The insertion of *Ds* into a locus of a gene results in a mutation at that locus. For example, the C locus in maize kernels makes a factor required for the synthesis of a purple pigment. Insertion of the *Ds* element in the locus

inactivates the gene, rendering the kernel colorless. This mutation is unstable, however. In the presence of the active *Ac* element, *Ds* is transposed away from the locus in some cells and the mutation reverts, giving rise to sectors of pigmented cells and thus to a purple-spotted kernel.

The Applicants have used a derivative of the *Ds* transposable element, *Ds2*, to produce a new mutant of the *Id* gene. This was accomplished by excision of *Ds2* (in the presence of active *Ac*) from a nearby gene on chromosome 1 and its subsequent insertion into the *Id* gene to produce *id**.

Through several generations of out-crosses and back-crosses, *id** was introduced into genetic backgrounds with or without active *Ac* elements. Data from these experiments show that *id*/id** plants with active *Ac* elements have a less severe phenotype than those with no *Ac* or *Id* plants; i.e., they exhibit fewer vegetative nodes and flower earlier. This result is expected if the *Ac* element mediates somatic excision of the *Ds2* element from the *id** allele during growth. Excision would restore *Id* function and result in partial restoration of normal development. Furthermore, the observation that these plants do not show patterns of defined sectoring (i.e., sharp demarcation of normal tissue juxtaposed to mutant tissue) suggests that *Id* acts non-cell-autonomously. This result implies that the *Id* gene product is either itself a diffusible factor, or that it regulates the production of a diffusible factor.

The above experiments, in which the effect of *Ac* on the flowering of *id** plants was studied, demonstrate that the flowering time of the maize plant can be regulated quantitatively by the amount of *id* gene product available. Wild-type (*Id*) plants from these families flowered at 9 to 11 weeks after planting. Plants homozygous for *id**, with no *Ac* present, had not flowered after 25 weeks, at which time the experiment was terminated due to frost. The

plants that were homozygous for *id** and which also had *Ac*, flowered anywhere from 15 to 22 weeks. Excisions of *Ds* occur in these plants due to the presence of *Ac*. These excisions restore *Id* function, and result in sufficient *Id* gene product to cause the plants to flower earlier than the plants with no *Ac*, but not sufficient *Id* gene product to cause them to flower as early as the wild type plants. The large range in flowering times presumably reflects the intrinsic variability in the timing and frequencies of *Ds* excisions from plant to plant. Fedoroff (1989), *supra*.

Another experiment examined the *Ac* effect on *id** plants more closely. The element *Ac* shows a "negative dosage" effect; that is, one copy of *Ac* causes many more *Ds* excisions than two or more copies of *Ac*. Fedoroff (1989), *supra*. The effect of *Ac* dosage on *id** plants was determined by planting seeds which were homozygous for *id** and which carried no *Ac*, one *Ac*, or two or more *Ac* elements per genome. If the amount of available *Id* product regulates flowering, then *Id** plants containing two or more *Ac* elements were expected to flower later than *id** plants with one *Ac* element but earlier than *id** plants with no *Ac* element. This experiment was performed under greenhouse conditions in which wild-type controls flowered after producing 12 to 13 leaves. None of the *id** plants lacking *Ac* elements flowered even after 24 leaves were produced. Of the *id** plants containing two or more *Ac* elements, 12.5% flowered after producing 21 to 23 leaves, whereas 87.5% of the plants did not flower even after producing 24 leaves. In contrast, 90% of the plants carrying one *Ac* element flowered after producing 16 to 24 leaves. The results demonstrate that *id** plants containing one *Ac* element (those with the greatest number of *Ds* excisions and therefore, the greatest amount of *Id* product) flower earlier than plants with more than one *Ac* element (although not as early as wild-type plants). The results also

suggest that varying the amount of functional *Id* gene product, *e.g.*, by varying the frequency of *Ds* excision through different doses of *Ac*, can induce a quantitative variation of the time of flowering.

Southern blot analysis using the *Ds2* element as a probe showed that a 4.2 kb *Sac*I fragment co-segregates with the *id** allele in more than 120 outcross progeny tested. This fragment is absent in plants that do not carry the *id** allele. Cosegregation of this fragment with the *id** allele is evidence that the gene is tagged with the *Ds2* transposon. This fragment was isolated by separation of *Sac*I cut genomic DNA on an agarose gel and excision of a region of the gel containing the fragment and sub-cloning into a plasmid vector to make a sub-library of genomic DNA in this region. The specific clone carrying the element was identified by probing the sub-library with the *Ds2* probe. From 60,000 clones analyzed, one was found to contain the 4.2 kb *Sac*I fragment. Restriction analysis showed that this recombinant clone carries a *Ds2* fragment flanked by maize DNA: 165 bp of DNA to one side of the *Ds2* element and 2.8 kb of DNA on the other side of the element (Figure 4). Southern blots of DNA from various plants using either of the flanking regions as probes showed that plants that are homozygous for the *id** allele contain a single *Sac*I band of 4.2 kb whereas those that contain only normal DNA have a single 2.9 kb *Sac*I fragment. Thus, the 4.2 kb fragment is the result of the insertion of the 1.3 kb *Ds2* element into the 2.9 kb *Sac*I fragment. Heterozygous plants contain both bands.

Further analysis of *id** and other *id* mutants has demonstrated that these mutants are variations of the normal *Id* gene which generally result from insertion or deletion of a genetic element at different sites within the *Id* gene sequence, or deletion of all or a part of the *Id* gene itself. DNA from mutant plants carrying the first *id*

allele to be identified, *id-R*, showed no hybridization to either of the flanking probes, indicating that this original allele is the result of a deletion of the *Id* gene. Another *id* allele, *id-Compeigne*, appears to have a 3 kb insertion into this fragment. These results provide convincing evidence that Applicants have tagged the *id* gene with *Ds2*.

Sequence analysis of the DNA immediately flanking the *Ds2* element of the *Id* gene revealed an open reading frame (ORF) into which the transposon has been inserted (Figure 4). When an RNA blot was probed with flanking DNA fragment that contained this ORF, a band of approximately 2.0-2.2 kb was evident in polyA+ RNA from apical meristem and, to a lesser extent, in mature leaf. Very little hybridization was detected in seedling RNA and none was detected in RNA from roots. This indicates that the ORF encodes a transcript and that the transcript is differentially expressed in specific plant tissues.

Analysis of the deduced amino acid sequence encoded by the ORF provided further evidence that this ORF is part of the *Id* gene and that it plays an important role in plant development. A comparison of this ORF to all proteins in current databases shows that it has significant homology to "zinc-finger"-like proteins identified in many different eukaryotes, including humans, mice, frogs (*Xenopus*) and *Drosophila* (Figure 4). Zinc-finger proteins are known as a class of diverse eukaryotic transcription factors that utilize zinc-containing DNA-binding domains and are important regulators of development. McKnight, S.L. and K.R. Yamamoto, eds. (1992) *Transcriptional Regulation.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, Vol. 1, p. 580. Zinc-finger proteins exert a regulatory function by mediating the transcription of other genes.

Results described herein show that the *Id* gene is

important in a crucial point in plant development (i.e., the transition from vegetative to reproductive growth) and that it functions by controlling the expression of other plant genes required for floral development. It is clearly a "switch" and nothing else in maize produces its effect (flower induction) without affecting the health and vigor of the plant. Conversely, mutation of *Id* alters or inhibits flower induction only; otherwise, the mutants are healthy and grow well.

Further evidence that the cloned DNA fragment is part of the *Id* gene was produced by generating two new alleles of *id* by imprecise excision of the *Ds2* element from the original *id** allele. Unlike *id**, these new alleles no longer respond to *Ac*; they are null mutants that appear not to flower at all. Sequence analysis shows that they have an altered sequence which results in a frame shift in the *Id* open reading frame caused by the excision of *Ds2* (Figure 7), and therefore, do not encode the same polypeptide as the *Id* gene.

Figure 7 illustrates the DNA and amino acid sequence of a portion of the normal *Id* ORF and its alteration as a consequence of *Ds* insertion and excision. The *id-Ds* mutation in *id* which is produced by insertion of the *Ds* transposon shows the 8 bp target site duplication (underlined) which is typical of *Ds* insertion. The null mutants, *id-X1* and *id-X2*, are stable, derivative alleles of *id* resulting from excision of *Ds2*. The *id-X1* allele has 7 bp of the duplication site remaining and an altered residue (T to A). The *id-X2* allele has 5 bp of the duplication site remaining with the same T to A transition as *id-X1*. The resulting amino acid residues show the frame shift in the ORF.

The entire clone carrying the 4.2 kb *Sac*I fragment was analyzed and the complete sequence of the genomic DNA flanking the *Ds2* element (SEQ ID NO:1) determined (Figure

2) using the information provided herein and methods of analysis known to those of ordinary skill in the field. The sequence of 2930 nucleotides comprises DNA of the *Id* gene. The deduced amino acid sequence (SEQ ID NO:2) encoded by this DNA is shown as sequence (c) in Figures 3A-3F.

Nucleotides 1 through at least 1890 (possibly through 2150) of SEQ ID NO:1 are transcribed. Nucleotides 176-1600 represent an intron. The approximate site of the *Ds2* transposon insertion is nucleotide 168. The ORF located between the *Nsi*I and *Sac*I restriction sites described *supra* (SEQ ID NO:3), is represented by nucleotides at positions 1-410 in Figures 3A-3F. (Note: the DNA strand shown in Figure 5 is complementary to that of Figures 3A-3F.)

The reproductive capacity of a plant directly affects its ability to yield seeds. Therefore, the ability to control flowering time is an important factor in the life cycle of the plant. The genetic studies of the *id* mutation of maize described herein indicate that the *Id* gene encodes a protein that is required for the transition to flowering. Through the use of transposon tagging, the Applicants have isolated and characterized the *Id* gene and, in particular, a portion of the zinc-finger regulatory region of this gene. Further, molecular analysis and comparison to eukaryotic animal regulatory proteins shows that the polypeptide encoded by this region is part of, if not the major component of, the regulatory *Id* protein that controls flower initiation and, very likely, also controls transition to reproduction from the vegetative growth stage of gymnosperms and lower plants, including the algae.

The DNA provided by this invention can be used to isolate homologous nucleic acids from other species of plants which encode regulatory genes for flowering similar in function to the *Id* gene. In this context, homology means an overall sequence identity of at least 50%,

preferably 70% or more for the zinc-finger portion of the *Id* allele. The identification and isolation of *Id*-type genes (homologues of *Id*) of other plant species is carried out according to standard methods and procedures known to those of ordinary skill in the art. See, *e.g.*, Sambrook, *et al.* (1989) *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. An example of this application is found in Example 5, *infra*.

By using these and other similar techniques, those of ordinary skill can readily isolate not only the *Id* gene in different cells and tissues of maize, but also homologues of the *Id* allele from other plant species. By example, *Id* genes in plants can be identified by preparing a genomic or cDNA library of a plant species; probing the genomic or cDNA library with all or a portion or a homologue of SEQ ID NO:1; identifying the hybridized sequences; and isolating the hybridized DNA to obtain the *Id* gene of that plant. Once identified, these genes can be restriction mapped, sequenced and cloned. In particular, the zinc-finger region or fragments thereof are especially effective as probes because of their conserved homology to other zinc-finger regions. Fragments as small as 20 bp in length can be used to hybridize to other zinc-finger regions.

By hybridization, it is meant that DNA and/or RNA are used in a hybridization analysis to detect complementary polynucleotides under conditions of moderate stringency according to methods described in Ausubel, *et al.* (1994) *Current Protocols in Molecular Biology*, Suppl. 26, John Wiley & Sons, Inc., New York, NY.

Other zinc-finger proteins that regulate phenomena other than flower initiation may be present in maize and other plants. Regulatory genes may control the germination of seeds, the height and shape of plants, the number of leaves, and the ripening of fruits to name a few possibilities. The isolation and characterization of these

genes as well as the genes responsible for initiation of the reproductive phase in plants would be of great significance and value in flower, food, and crop production in general. Such zinc-finger genes in plants can be identified by preparing a genomic or a cDNA library of a plant species; probing the genomic or cDNA library with all or a portion or a homologue of the *Id* gene, described herein, such as SEQ ID NO:1, under conditions appropriate for hybridization of complementary DNA identifying the hybridized DNA; and isolating the hybridized DNA to obtain the zinc-finger gene in that plant. The zinc-finger genes can then be restriction mapped, sequenced and cloned.

This invention also provides nucleic acids and polypeptides with structures that have been altered by different means, including but not limited to, alterations using transposons, site-specific and random mutagenesis, and engineered nucleotide substitution, deletion, or addition.

A transposon method of producing an allele of the *Id* gene with an altered function in a plant can comprise: inserting the *Ds* transposon or another nonautonomous transposable element into the *Id* gene, and then excising the *Ds* transposon with the *Ac* transposon or another autonomous transposable element to produce an altered *Id* allele in the plant.

A further example of a method of producing an allele of the *Id* gene with an altered function in a plant comprises altering the molecular structure of the *Id* gene *in vitro* using molecular genetic techniques (e.g., site specific mutagenesis), and then inserting the altered *Id* gene into a plant to produce an altered *Id* allele in the plant.

These techniques can give rise to *Id* homologs which demonstrate dramatically different functions from the corresponding naturally-occurring protein. For example,

site-directed mutagenesis can be used to produce *Id* alleles that encode specific substitutions of amino acid residues and it can then be determined what amino acids are required to produce a functional gene, the product of which induces a reproductive response in plants. Likewise, *Id* alleles can be engineered to produce proteins that have novel functions, such as flower induction earlier than that of the naturally-occurring plant.

There are many varieties of maize that have evolved a wide range of flowering times depending on the environmental conditions in which they are grown. In particular, day-length (as dictated by latitude) determines when a plant will flower. The *Id* gene is a determinant of flowering time in all of these maize variants, and flowering time may be correlated to specific variations in the *Id* gene product. In fact, the *Id* gene may be the major determinant of floral evocation.

The *Id* gene or a homologue thereof can be altered and introduced into a maize plant to alter the flowering time of a particular type of maize so that it can be grown in a different latitude from the one in which the parent strain was developed. Thus, an engineered *Id* gene can be incorporated into a maize line that has been bred for other traits (e.g., high yield and disease resistance), to produce a maize line that can be grown at many different latitudes. Lowering the level of Id protein using antisense constructs or co-suppression (see below) can delay flowering time, while increasing the level of *Id* by overexpression or through earlier production (*Id* gene coupled to a different promoter) of the protein can induce plants to flower sooner. Further, putting the sense or antisense *Id* gene under the control of different inducible promoters can permit flowering time to be controlled when subjected to specific environmental conditions or to applied chemicals.

Co-suppression refers to the overexpression of an endogenous or an introduced gene (transgene) wherein the extra copies of the gene result in the coordinate silencing of the endogenous gene as well as the transgene, thus reducing or eliminating expression of the trait. See, for example, Jorgensen *et al.*, U.S. Patents No. 5,034,323 and No. 5,283,184. The transgene is introduced in a sense orientation and does not require a full length sequence or a lute homology to the endogenous sequence intended to be re ssed.

Expression f the endogenous gene may also be suppressed thrc gh the integration of an oligonucleotide having an identical or homologous sequence to that of the DNA strand complementary to the strand transcribing the endogenous gene. Antisense oligonucleotides comprise a specific sequence of nucleotides that provide an RNA which stably binds to the RNA transcribed from the endogenous gene, thus preventing translation. See, Shewmaker *et al.*, U.S. Patent No. 5,107,065.

Other oligonucleotides of this invention called "ribozymes" can be used to inhibit or prevent flowering. Unlike antisense and other oligonucleotides which bind to an RNA, a DNA, or a protein, ribozymes are catalytic RNA molecules which can bind and specifically cleave a target RNA, such as the transcription product of an endogenous *Id* gene. Ribozymes designed to cleave at specific sites can inactivate such an RNA molecule. Thus reduction of an *Id* product can be achieved by introduction of DNA which encodes a ribozyme designed to specifically cleave transcripts of endogenous *Id* genes in an endonucleolytic manner.

Of the known classes of ribozymes, the group I intron and hammerhead ribozymes are useful candidates to convert for targeted cleavage of an *Id* transcript since they have short (4-12 base) recognition sequences; however, other

types of ribozymes can be developed for site-specific cleavage of *Id* mRNA. See, Cech, T.R. (1988) *J. Amer. Med. Assoc. 260*:3030-3034.

The above strategies to delay or completely abolish flowering depend upon the use of antisense and similar technologies. An alternative strategy can be devised based upon the use of "dominant-negative" mutant proteins. Certain types of mutations can be introduced into regulatory proteins that render them non-functional, but permit the mutant proteins to compete with the wild-type proteins for their targets. Such competition by a non-functional protein means that overexpression of the mutant protein can be used to suppress the activity of the wild-type protein. Dominant-negative mutations of zinc-finger transcription factors have been constructed in fruit-flies and in human cells by deleting the activation/silencer domain while retaining the DNA-binding zinc-finger domain. The over-expressed mutant protein then competes out the wild-type protein by binding non-productively to the DNA targets. O'Neill, E.M. *et al.* (1995) *Proc. Nat'l. Acad. Sci. USA 92*: 6557-6561. In plants, dominant-negative strategies have been used successfully with other types of regulatory proteins. See, Boylan, M. *et al.* (1994) *Plant Cell 6*: 449-460; Rieping, M. *et al.* (1994) *Plant Cell 6*: 1087-1098; and Hemerly, A. *et al.* (1995) *EMBO J. 14*: 3925-3936.

A dominant-negative mutant of the *Id* protein can be constructed by using a truncated version of the *Id* gene that contains only the sequences encoding the zinc-finger domain (the presumptive DNA-binding domain), and is missing the activation domain. If this truncated gene is introduced into maize plants under the control of a strong promoter, the result will be maize plants that are either severely delayed in flowering or are unable to flower. Therefore, the truncated dominant-negative *Id* gene can be

substituted for the antisense *Id* gene in all of the constructs used to delay flowering herewith described.

The dominant-negative *Id* gene approach has an advantage over the antisense construct when engineering delayed flowering into crops other than maize. The antisense strategy depends on initially cloning part or all of the *Id* gene from each crop species, then expressing these genes in an inverted orientation. Antisense suppression depends on expression of the complementary nucleotide sequences, which will vary from one crop species to another. In contrast, the dominant-negative strategy depends only upon the functional conservation of the protein and its target sites. Overall, this is a much less stringent requirement than nucleotide sequence conservation. Several known examples of regulatory genes encoding transcription factors perform similar functions when expressed in widely divergent species of plants. See, *e.g.*, Lloyd, A.M. *et al.* (1992) *Science 258*: 1773-1775; Irish, V.F. and Y.T. Yamamoto (1995) *Plant Cell* 7:1635-1644. This type of functional conservation implies that the dominant-negative version of the maize *Id* gene can work similarly in other crop species as well. It can certainly be expected to function in other cereal species and perhaps in all monocotyledonous plants.

For application to dicots, it could be advantageous to first isolate a more closely-related *Id* homolog from a dicotyledonous species (*e.g.*, tobacco or *Arabidopsis*), and construct a dominant-negative derivative as described above (by removing all sequences other than the zinc-finger DNA binding domain). This dicot version of dominant-negative *Id* can then be used for all dicot plants. Thus, application of dominant-negative technology to a wide range of crops can be achieved without the need to clone *Id* genes from every crop.

Any suitable technique can be used to introduce the

nucleic acids and constructs of this invention to produce transgenic plants with an altered floral induction time. For grasses such as maize, microprojectile bombardment (see for example, Sanford, J.C., *et al.*, U.S. Patent No. 5,100,792 (1992) can be used. In this embodiment, a nucleotide construct or a vector containing the construct is coated onto small particles which are then introduced into the targeted tissue (cells) via high velocity ballistic penetration. The vector can be any vector which expresses the exogenous DNA in plant cells into which the vector is introduced. The transformed cells are then cultivated under conditions appropriate for the regeneration of plants, resulting in production of transgenic plants. Transgenic plants carrying the construct are examined for the desired phenotype using a variety of methods including but not limited to an appropriate phenotypic marker, such as antibiotic resistance or herbicide resistance, or visual observation of the time of floral induction compared to naturally-occurring plants.

Other known methods include Agrobacterium-mediated transformation (see for example Smith, R.H., *et al.*, U.S. Patent No. 5,164,310 (1992), electroporation (see for example, Calvin, N., U.S. Patent No. 5,098,843 (1992)), introduction using laser beams (see for example, Kasuya, T., *et al.*, U.S. Patent No. 5,013,660 (1991)) or introduction using agents such as polyethylene glycol (see for example Golds, T., *et al.* (1993) *Biotechnology*, 11:95-97), and the like. In general, plant cells may be transformed with a variety of vectors, such as viral, episomal vectors, Ti plasmid vectors and the like, in accordance with well known procedures. The method of introduction of the nucleic acid into the plant cell is not critical to this invention.

The transcriptional initiation region may provide for

constitutive expression or regulated expression. Many promoters are available which are functional in plants. Illustrative promoters include the octopine synthase promoter, the nopaline synthase promoter, the cauliflower mosaic virus (35S) promoter, the figwort mosaic virus (FMV) promoter, heat-shock promoters, ribulose-1,6-biphosphate (RUBP) carboxylase small subunit (ssu), tissue specific promoters, and the like. The regulatory region may be responsive to a physical stimulus, such as light, as with the RUBP carboxylase ssu, differentiation signals, or metabolites. The time and level of expression of the sense or antisense orientation can have a definite effect on the phenotype produced. Therefore, the promoters chosen, coupled with the orientation of the exogenous DNA, will determine the effect of the introduced gene.

Transgenic plants of this invention can contain an exogenous nucleic acid which alters the time of floral induction so that floral induction is earlier than that of a plant of the same variety without said exogenous nucleic acid when grown under identical conditions. Alternatively, transgenic plants containing an exogenous nucleic acid which alters the time of floral induction so that floral induction is delayed or inhibited compared to floral induction in a plant of the same variety without said exogenous nucleic acids when grown under identical conditions.

Further, this invention includes a method of producing a transgenic plant having an altered time of flower induction, comprising introducing into plant cells an exogenous nucleic acid whose presence in a plant results in altered time of induction of flower development, and maintaining plant cells containing the exogenous nucleic acid under conditions appropriate for growth of the plant cells, whereby a plant having an altered reproduction induction time is produced. Organisms to which this method

can be applied include: angiosperms (monocots and dicots), gymnosperms, spore-bearing or vegetatively-reproducing plants and the algae.

Transgenic plants containing the *Id* recombinant constructs can be regenerated from transformed cells, tissues or plant parts by methods known to those of skill in the art. Plant part is meant to include any portion of a plant capable of producing a regenerated plant. Thus, this invention encompasses a cell or cells, tissue (especially meristematic and/or embryonic tissue), protoplasts, epicotyls, hypocotyls, cotyledons, cotyledonary nodes, pollen, ovules, stems, roots, leaves, and the like. Plants may also be regenerated from explants. Methods will vary according to the plant species.

Seed can be obtained from the regenerated plant or from a cross between the regenerated plant and a suitable plant of the same species. Alternatively, the plant may be vegetatively propagated by culturing plant parts under conditions suitable for the regeneration of such plant parts.

Isolated and purified *Id* or *id* protein or polypeptides, and epitopic fragments thereof, can be used to prepare antibodies for localization of sites of *Id* regulation and to analyze developmental pathways in plants. For example, antibodies that specifically bind an *Id* protein can be used to determine if and when the protein is expressed in specific cells or tissues of the plant. This information can be used to determine how *Id* acts to induce flowering and to alter flower induction pathways.

Antibodies of the invention can be polyclonal, monoclonal, or antibody fragments, and the term antibody is intended to encompass polyclonal antibodies, monoclonal antibodies and antibody fragments. Antibodies of this invention can be raised against isolated or recombinant *Id*

or *id* proteins or polypeptides. Preparation of immunizing antigen, and antibody production can be performed using any suitable technique. A variety of methods have been described (see e.g., Harlow, E. and D. Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; Ausubel *et al.* (1994) *Current Protocols in Molecular Biology*, Vol. 2, Chapter 11 (Suppl. 27) John Wiley & Sons: New York, NY).

Antibodies of this invention can be labeled or a second antibody that binds to the first antibody can be labeled by some physical or chemical means. The label may be an enzyme which is assayed by the addition of a substrate which upon reaction releases an ultraviolet or visible light-absorbing product or it can be a radioactive substance, a chromophore, or a fluorochrome. E. Harlow and D. Lane (1988) *supra*.

Isolated polypeptides of this invention can also be used to detect and analyze protein/protein interactions. Fusion proteins for this purpose can be prepared by fusing *Id* DNA encoding a functional *Id* polypeptide with heterologous DNA encoding a different polypeptide (one not related or homologous to the *Id* polypeptide), such as a protein tag. The resulting fusion protein can be prepared in a prokaryotic cell (e.g. *E. coli*), isolated, labeled and used essentially like antibodies to detect binding sites of *Id* alleles and *Id*/protein interactions. See Ron and Dressler (1992) *Biotech* 13:866-69; Smith and Johnson (1988) *Gene* 67:31-40.

Maize lines that are adapted to temperate latitudes flower prematurely when planted in the tropics due to the shorter daylengths. The premature flowering results in severely reduced yields. Salamini, F. (1985) *Breeding Strategies for Maize Production Improvement in the Tropics*. Brandolini, A. and Salamini, F., eds. Food and Agriculture Organization of U.N., Istituto Agronomico Per L'Oltremare,

Firenze, Italy. One of skill in the art will recognize that the cloned *Id* gene can be used to overcome this problem. Transgenic maize plants can be generated in which the *Id* gene is inserted in the antisense orientation under the control of a weak promoter (Figure 8A). The weak promoter used should be constitutively active during development, at least in the shoot meristem. Since *Id* appears to be non cell-autonomous, exact specification of the site of action of the promoter is not necessary. An example of a weak promoter useful for this application is the nopaline synthase (*nos*) promoter, from T-DNA, shown to be weakly constitutive in maize. Callis, *et al.* (1987) *Genes Dev.* 1:1183-1200. Another is a cyclin promoter from maize. Cyclins are cell division proteins found in plants, animals and yeasts. Plant cyclin transcripts are expressed in meristems and tissues with proliferating cells at low levels, but are not expressed elsewhere. Renaudin, *et al.* (1994) *PNAS* 91:7375-7379. The cyclin promoters are easily isolated by using Applicants' full-length cDNA clones for cyclin 1b or cyclin III as probes, to pull out the flanking upstream genomic sequences from a maize genomic library using standard isolation and cloning techniques. See, Sambrook, *et al.*, *supra*; Freeling and Walbot, *supra*. Those skilled in the art will recognize the other weak promoters intended to be encompassed by the invention that have the characteristics necessary to carry out this embodiment of the invention.

An example of a construct useful for the above application is illustrated in Figure 8A. The cDNA for the *Id* gene is ligated downstream from the promoter, in the antisense orientation. The ADH1 intron is required for RNA stability, and the 3' end of the *nos* gene is added to ensure efficient polyadenylation. Callis, *et al.* (1987) *supra*. The DNA is introduced into maize plants by standard methods such as those described above, using the *bar* gene

for resistance to the herbicide Basta as the transformation marker. Gordon-Kamm, *et al.* (1990) *Plant Cell* 2:603-618; Freeling and Walbot (1993) *supra*.

Any construct or vector which expresses the exogenous DNA in plant cells into which it is introduced can be used, such as the pMON530 vector carrying the 35S promoter. Another useful vector or construct of the present invention is exogenous DNA encoding the *Id* protein inserted in the antisense orientation into the pMON530 vector downstream of a weak promoter to delay flowering in an early-flowering variety.

Similar constructs can be used for other cereals, e.g., rice, barley, and other monocotyledonous crops. For antisense applications, it may be necessary to first isolate the homologous cDNA from the species to be modified. It will be recognized that the maize *Id* clone can be used as a probe for this purpose, screening for *Id* homologues from cDNA libraries of the other cereal species. The *Id* homologue for the species to be engineered can then be inserted as a substitution for the maize *Id* gene in the constructs of Figure 8A.

The same technique can be extended to dicotyledonous plants as well. Delaying flowering time for some of these crops can result in advantages similar to those cited for maize, *i.e.*, a longer vegetative growth period that results in higher yields of fruits and seeds. Gottschalk and Wolff (1983) *Induced Mutations in Plant Breeding*, Springer-Verlag, Berlin, Heidelberg. In addition, some dicotyledonous plants are valuable chiefly for the products of vegetative growth (e.g., spinach, tobacco, etc.), and, in these plants, extended vegetative growth will result in higher and more efficient yields of products. Antisense constructs can be designed using *Id* homologues isolated from these species, as shown in Figure 8B, and transgenic plants generated by T-DNA transformation, preferably using

Agrobacterium transformation techniques, but also by other standard techniques. Lycett, G.W. and D. Grierson (1990) *Genetic Engineering of Crop Plants*, Butterworths, London; Setlow, J.K. (1994) *Genetic Engineering Principles and Methods*, Vol. 16, Plenum Press, New York.

Maize varieties that are adapted to tropical latitudes flower extremely late when grown in temperate latitudes (Salamini, *supra*), reaching heights of 15-20 feet, with 30 leaves at flowering (compared to about 20 leaves on the average temperate variety). This is not only inconvenient for handling and harvesting, but makes the plants vulnerable to late season frost damage. A strategy to induce earlier flowering in these plants is to express the cloned *Id* gene early in the vegetative development of these varieties by inserting the gene in the sense orientation under a constitutive promoter (Figure 9A). A strong or weak promoter can be used, such as the CaMV 35S (strong) promoter or the *nos* (weak) promoter, both of which function in maize. Callis, *et al.* (1987) *supra*. The constructs and transformation methods for this purpose are similar to those used in the antisense application described above except for the orientation of the *Id* gene.

It will be recognized that this technique can be adapted for other cereal species and for monocots, in general, using the same constructs or constructs that are similar in principle. In fact, homologues of *Id* may not be required for early expression because a maize *Id* gene product could function adequately in other monocotyledons, including cereals, to promote earlier flowering.

In another embodiment of this invention, earlier flowering of dicotyledonous plants can be provided by transforming target plants or plant cells with the maize *Id* gene product or an *Id* homologue. Because maize genes have been demonstrated to function efficiently in dicots, it may not be necessary to isolate the homologous gene from the

species to be transformed. For example, the maize *R* and *C* genes function in the dicot *Arabidopsis* when expressed under control of the CaMV 35S promoter. Lloyd, *et al.* (1992) *Science* 258:1773-1775. The construct delineated in Figure 9B can be used for expression of an *Id* gene or homologue in a dicot, and can be inserted with T-DNA transformation or other standard techniques such as those already described.

Drought stress can cause severe reduction in yields due to damage to the plant. In addition, the flowering time can be affected. Many plants flower prematurely when stressed. In maize, drought stress can result in the tassel developing much earlier than the ear, resulting in reduced yields or no yields. Some of these problems can be alleviated if the overall flowering time of the plant was delayed during a period of drought. This delay would allow the plant to grow vegetatively for a longer period of time than normal, so that it can recover from drought damage before it flowers. The *Id* gene can be used for this purpose, if it is introduced into the plants in the antisense orientation as described earlier, but combined with a drought-inducible promoter instead of a constitutive promoter. Any drought-inducible promoter can be used. For example, a promoter for the RAB-17 gene, which is induced by drought as well as other stresses, presumably as a result of its regulation by the plant hormone ABA can be used. Vilardell, *et al.* (1990) *Plant Mol. Biol.* 14:423-432. A second type of promoter which can be used is the maize *hsp70* heat shock promoter, which is induced in response to high temperatures 37° to 42°C. Callis, *et al.* (1988) *Plant Physiol.* 88:965-968.

A useful vector or construct to produce plants responsive to environmental effects is produced by inserting the exogenous DNA encoding the *Id* protein in the antisense direction into the pMON530 vector downstream of a

drought-induced promoter to delay flowering in response to drought. Several constructs for this purpose are illustrated in Figure 10A.

Again, this technique can be extended to monocots in general, including other cereals, with the same constructs as in Figure 10A or a similar construct, but using the homologue of the *Id* gene for the particular cereal being transformed if necessary.

The extension of this technique to dicotyledonous crops can be performed using appropriate drought inducible promoters that function in dicotyledonous plants. The promoter of the *Arabidopsis Atmyb2* can be used as a general ABA-responsive, drought and stress-induced promoter. Urao, *et al.* (1993) *Plant Cell* 5:1529-1539. The soybean heat-shock promoter can also be used. Schoffl, *et al.* (1989) *Mol. Gen. Genet.* 217:246-253. Constructs including such promoters are illustrated in Figure 10B. Since this application depends upon antisense expression, it may be necessary to use the homologue of the *Id* gene from the crop species that is being engineered, rather than the maize *Id* gene.

Of particular use are maize plants in which flowering is completely absent; *i.e.*, knocked out. Maize plants that do not flower will continue to grow vegetatively, producing a large biomass which can be harvested for silage purposes. However, if the *Id* gene is knocked out completely for the purposes of producing silage, the transgenic plants will never flower and no hybrid seeds can be produced.

One method of this invention for generating hybrid seeds of transgenic corn is to produce transgenic plants with the *Id* gene in the antisense orientation, but under the control of a regulatory sequence called the GAL4 binding site. As a consequence, the antisense *Id* gene is not expressed unless the GAL4 protein is present. GAL4 is a transcription factor from yeast, which has been

demonstrated to work in plants such as tobacco (Ma, J., *et al.* (1988) *Nature 334*:631-633), as well as in corn (McCarty, D. *et al.* (1991) *Cell 66*:895-905. It activates transcription of genes which contain the GAL4 binding site in the promoter.

In this embodiment, a transgenic inbred containing the silent antisense *Id* gene and the GAL4 binding site is crossed to another transgenic inbred which expresses the GAL4 gene constitutively, either under a weak promoter (to delay flowering for growth of corn in lower latitudes), or under a strong promoter (to abolish flowering for silage production). Each inbred flowers normally. However, the hybrid expresses the antisense *Id*, and flowering is delayed or absent, depending upon the promoter used to drive the GAL4 gene. A similar modification can be made for other plants, either monocots or dicots, using the appropriate *Id* homolog.

Constructs using the GAL4 binding site are illustrated in Figures 11A, 11B, 11C and 11D. Thus, in maize, an inbred comprising the construct illustrated in Figure 11A is crossed with an inbred comprising the construct of Figure 11C. Flowering is delayed in the resulting hybrid when the GAL4 gene is under the control of CaMV 35S (P35s). When the GAL4 gene is under the control of the *nos* (Pnos) or *cyclin* (Pcyclin) promoters, however, flowering is only delayed in the hybrid. In dicots, similar results are obtained by crossing the plant comprising the construct shown in Figure 11C to the plant comprising the construct shown in Figure 11D.

The applications described above illustrate the use of antisense *Id* constructs. It will be recognized by those of skill in the art that any suitable construct, for example, the dominant-negative version of the *Id* gene, can be substituted for the antisense constructs to practice the methods of this invention.

Although the *Id* gene was isolated from maize, it is likely that homologues of *Id* exist in other grain crops, and most likely in all other plants. Applicants have initial evidence that a close relative of *Id*, as determined by sequence homology, exists in dicotyledonous plants as well. If these homologues in other species are also important to the control of flowering time, then the manipulation of flowering time of many agriculturally important crops would be possible. Using the compositions and methods described herein, a skilled artisan can use known procedures to alter initiation of the reproductive phase of other grains such as sorghum, rye, wheat, etc., as well as in other commercially important plants.

For example, modifications of flowering time can be used to affect the time of ripening of fruit, time of production of flowers, size and quality of seed, latitude at which varieties can be grown, and the like. Flowering time may be modulated so that flowering is initiated at different times on different parts of the same plant. This invention also provides a means to eliminate the need for detasseling in the production of maize and sorghum hybrids. Although it appears that *Id* does not act in a cell autonomous manner, it may be that the *Id* signal is localized to certain areas of the plant and thus *Id* must be transcribed or *Id* mRNA activated in several areas of the plant to induce flower development in each of these areas. Corn and sorghum both produce male flower organs (tassels) at the top (apex) of the plant. Female flower organs are produced on lower portions in the axils. Through the use of tissue-specific or other selective promoters coupled to the *Id* gene, it is possible to inhibit or prevent the production of pollen in the apex of the plant while selectively inducing reproductive development of the female reproductive organs on other parts of the plant. Or, after normal flower induction, development of male reproductive

organs can be inhibited or pollen-producing tissues or cells can be induced to revert to vegetative phase by coupling *Id* antisense production to the formation of cells specific to pollen production (such as tapetal cells).

Another application of this technology is to increase the vegetative phase (and therefore increase the number of leaves produced) of crops that are harvested as leaves (e.g., lettuce, cabbage, spinach, maize) and thereby increase yield of these crops by delaying flowering. Thus, any plant may be employed in accordance with this invention, including angiosperms, gymnosperms, monocotyledons, and dicotyledons. Plants of interest include cereals such as wheat, barley, maize, sorghum, triticale, etc.; other commercially-valuable crops, such as sunflower, soybeans, safflower, canola, etc.; fruits, such as apricots, oranges, apples, avocados, etc; vegetables, such as carrots, lettuce, tomatoes, broccoli, etc; woody species, such as poplar, pine, oak, etc; and ornamental flowers, such as clematis, roses, chrysanthemums, tulips, etc.

The following examples describe specific aspects of the invention to illustrate the invention and provide a description of the methods used to isolate and identify the *Id* gene. The examples should not be construed as limiting the invention in any way.

All citations in this application to materials and methods are hereby incorporated by reference.

Example 1

Transposon Tagging:

Plants were grown under normal field conditions at Uplands Farm Agricultural Field Station, Cold Spring Harbor Laboratory, during the summers of 1989 through 1994. Standard maize genetic techniques were used in all crosses and analytical procedures. Freeling, M. and Walbot, V.

(1993) *The Maize Handbook*. Springer-Verlag, New York; Gottschalk, W. and Wolff, G. (1983) *Induced Mutations in Plant Breeding*. Springer-Verlag, Berlin Heidelberg.?

The *Id* gene maps near the kernel pigmentation gene, *Bz2*, on chromosome 1. A mutable allele of the *Bz2* gene, *bz2-m2*, is the result of an insertion of a *Ds2* transposon at this locus. Dooner, *et al.* (1985) *Mol. Gen. Genetics* 200:240-246. (*Ds2* is a defective derivative of the *Ac/Ds* family of transposable elements and is able to transpose only in the presence of an *Ac* element which provides transposase.) Taking advantage of the proximity of *Id* to *bz2-m2*, and the fact that *Ac/Ds* elements transpose preferentially to linked sites, Applicants selected for *id* mutants from germinal revertants in the *bz2-m2* population; i.e., by selecting for completely purple kernels that resulted from germinal excision of the *Ds2* element (i.e., *bz2-m2* to *Bz2*), an F1 population with the Ds2 element inserted elsewhere was generated. From an F2 population of these revertants one *id* mutant was isolated from 600 families examined and designated *id**. Crosses of *id** to known alleles of *id* (*id-R*, for example) confirmed that it is allelic to the *id* mutation on chromosome 1.

Example 2

DNA Analysis:

All molecular biological procedures were performed essentially as described in Sambrook, J., *et al.* (1989) *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. Methods for the analysis of maize DNA and RNA were done according to Freeling, M. and Walbot, V. (1993) *supra*.

For Southern blot analysis, 2-4 mg of maize DNA extracted from leaves was restricted with *Sac*I and electrophoresed on a 1% agarose gel prior to transfer onto Nitrocellulose membranes. For Ds2 probing, an internal 108

bp fragment of the Ds2 transposon was isolated from a plasmid carrying this portion of Ds2 and cut with restriction enzymes *Bam*HI and *Eco*RI. This fragment was purified from a low melting point agarose gel and radioisotope-containing nucleotides (32P-dATP and 32P-dGTP) were incorporated into the fragment by random primed labeling using a kit from Boehringer-Mannheim. The labeled fragment was used to probe Southern blots using standard formamide hybridization solutions containing 10% dextran sulfate.

To isolate the Ds2-hybridizing 4.2 kb *Sac*I fragment, 100 mg of DNA from a single plant was digested with *Sac*I and electrophoresed on a 1% low-melting agarose gel. A region of the gel between 4 and 5 kb, marked by side markers, was excised from the gel and the DNA contained within the fragment was purified. The DNA was ligated (T4 DNA Ligase, New England Biolabs) into the plasmid vector pLITMUS29 (New England Biolabs) that had been cut with *Sac*I and phosphatase treated (Shrimp Alkaline Phosphatase, U.S. Biochemical) to remove 5' phosphate groups to prevent self ligation. Recombinant plasmids were transformed into the *E. coli* DH10B cells by electroporation and plated on L-agar plates containing 100 μg/ml ampicillin. Approximately 60,000 ampicillin-resistant colonies were grown up on plates and then replica transferred to nitrocellulose membranes. Colonies on filters were lysed and their DNA fixed to the membrane. To determine which colonies carried a recombinant plasmid that hybridized to the Ds2 element, the filters were probed with a labeled Ds2 fragment probe. Hake, *et al.* (1989) *EMBO J.*, 8:15-22. One colony from 60,000 screened was found to have a plasmid that had a Ds2 element. Restriction analysis of this recombinant plasmid revealed approximately 2.9 kb of genomic DNA to one side of the 1.3 kb Ds2 element and 165 bp on the other side. Sequence analysis of a portion of the flanking DNA was

performed by using primers that hybridized to sequence within the plasmid vector and within the Ds2 element itself. The dideoxy chain termination sequencing method was used to sequence double-stranded plasmid DNA.

Example 3

RNA Analysis:

Northern blot analysis of polyA RNA from various maize tissues was performed using the 165 bp genomic DNA region to the right flank of the Ds2 element as a probe. RNA was extracted from apical meristem tissue, young and old leaf tissue and from root tips, and 1 μg of each poly A+ mRNA from each sample electrophoresed on a 1.1% agarose gel containing formaldehyde and then transferred to Genescreen nylon membranes. The 165 bp fragment was labeled as described above, and used to probe the blots.

Example 4

Determination of the *Id* Gene Sequence from the Isolated Genomic Clone:

The genomic clone was sequenced by the dideoxy method as described in Sambrook, *et al.*, *supra*. The strategy used was called "primer walking". Oligonucleotide primers which hybridize to the plasmid vector were used to obtain initial sequence data for the ends of the fragment. This sequence data was then used to synthesize new primers within the sequenced region, which enabled further sequencing into the genomic clone in a reiterative process until the entire fragment was sequenced. Approximately 200 to 350 bp of sequence was read from each primer.

Example 5

Identification and Isolation of Regulatory Genes from other Plant Species:

To identify and isolate regulatory genes in other

species of plants which are homologous to the *Id* gene, the DNA sequence encoding the *Id* ORF or another fragment of the *Id* gene is used as a probe to screen plant cDNA libraries made of mRNA derived from tissues which express regulatory genes (Sambrook, *et al.* (1989) *supra*; Freeling and Walbot (1993) *supra*). cDNA libraries constructed from mRNA derived from seedlings and from immature inflorescence tissue are especially likely to contain these genes. Similar libraries from maize have been used successfully by Applicants to obtain cDNA clones of maize cell division cycles genes, such as *cdc2* (Colasanti, *et al.* (1991) *PNAS*, 88:3377-3381) and the cyclins (Renaudin, *et al.* (1994) *PNAS*, 91:7375-7379) by using short DNA probes for these genes. Clones which hybridize with the radioactive probes are identified and isolated, and a sequence analysis performed by standard methods as described in Sambrook, *et al.*, *supra*.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described specifically herein. Such equivalents are intended to be encompassed in the scope of the following claims.

Claims

We claim:

1. Isolated DNA which is an *Id* gene or a portion thereof.

2. Isolated RNA or a portion thereof encoded by the DNA of Claim 1.

3. Isolated *Id* polypeptide or portion thereof.

4. Isolated DNA complementary to an *Id* gene or portion thereof.

5. Isolated DNA of a plant that hybridizes under moderate stringency conditions to the DNA of Claim 1 or that shows at least 50% homology to the zinc finger region of the *Id* gene.

6. Isolated DNA comprising an *Id* gene, or a portion thereof, the *Id* gene comprising all or a portion of SEQ ID NO:1, or homologous DNA.

7. An *Id* gene encoding a polypeptide comprising SEQ ID NO:2.

8. A polypeptide or portion thereof encoded by the DNA according to Claim 6.

9. A plant or plant part which contains:
   a) an exogenous or altered *Id* or *Id*-type gene; or
   b) DNA comprising an *id** gene;
   c) DNA comprising an *Id* antisense construct;
   d) DNA encoding a dominant-negative mutant *Id* protein.

10. A seed of a plant of Claim 9.

11. A tissue culture of the plant or a plant part of Claim 9.

12. A plant or plant part according to Claim 9 wherein the plant is maize or sorghum or the plant part is derived from maize or sorghum.

13. The seed according to Claim 10 wherein the seed is a maize or sorghum seed.

14. A tissue culture according to Claim 11 wherein the tissue is maize or sorghum tissue.

15. A transgenic plant, transgenic plant part or transgenic plant cell containing exogenous DNA that alters the time of flower induction.

16. The plant or plant part according to Claim 15 wherein the plant is maize or sorghum or the plant part or plant cell is derived from maize or sorghum.

17. A transgenic plant containing an exogenous nucleic acid which alters the time of floral induction so that floral induction is earlier than that of a plant of the same variety without said exogenous nucleic acid when grown under identical conditions.

18. A transgenic plant containing an exogenous nucleic acid which alters the time of floral induction so that floral induction is delayed or inhibited compared to floral induction in a plant of the same variety without said exogenous nucleic acid when grown under identical conditions.

19. A method of producing a transgenic plant having an altered time of flower induction, comprising introducing into plant cells an exogenous nucleic acid whose presence in a plant results in altered time of induction of flower development, and maintaining plant cells containing the exogenous nucleic acid under conditions appropriate for growth of the plant cells, whereby a plant having an altered reproduction induction time is produced.

20. The method of Claim 19 wherein the transgenic plant is selected from the group consisting of: angiosperms, gymnosperms, monocots and dicots.

21. The method of Claim 19 wherein the exogenous nucleic acid is all or a portion of the *Id* gene or a homologue thereof.

22. The method of Claim 19 wherein the exogenous nucleic acid is all or a portion of the *id** gene or a homologue thereof.

23. A method of identifying an *Id* gene in a plant, comprising the steps of:

a) preparing a genomic DNA library or a cDNA library of a plant;

b) probing said genomic DNA library or cDNA library with all or a portion of or a homologue of the *Id* gene described in part by SEQ ID NO:1;

c) identifying the hybridized DNA; and

d) cloning the hybridized DNA to obtain the *Id* gene.

24. A method of identifying a gene encoding a zinc-finger protein in a plant, comprising the steps of:

a) preparing a genomic DNA library or a cDNA library of a plant;

b) probing said genomic DNA library or cDNA library with all or a portion or a homologue of the *Id* gene;

c) identifying the hybridized DNA; and

d) sequencing the hybridized DNA to obtain the gene encoding a zinc-finger protein.

25. The method of Claim 24 wherein the portion or homologue of the *Id* gene encodes the zinc-finger portion of the *Id* protein.

26. A method of producing an allele of the *Id* gene with an altered function in a plant comprising:

a) inserting a nonautonomous transposable element into the *Id* gene; and

b) excising the nonautonomous transposable element by use of an autonomous transposable element, to produce an altered *Id* allele in the plant.

27. A method of producing an allele of the *Id* gene with an altered function in a plant comprising:

a) altering the molecular structure of the *Id* gene *in vitro* using molecular genetic techniques; and

b) inserting the altered *Id* gene into a plant to produce an altered *Id* allele in the plant.

28. An antibody or antibody fragment which binds a polypeptide comprising SEQ ID NO:2, or a portion thereof.

29. An *Id* fusion protein comprising all or a portion of SEQ ID NO:2 or a homologue, and a polypeptide which is not related or homologous to SEQ ID NO:2.

30. A ribozyme which cleaves and inactivates the RNA transcript of an *Id* gene or its functional equivalent.

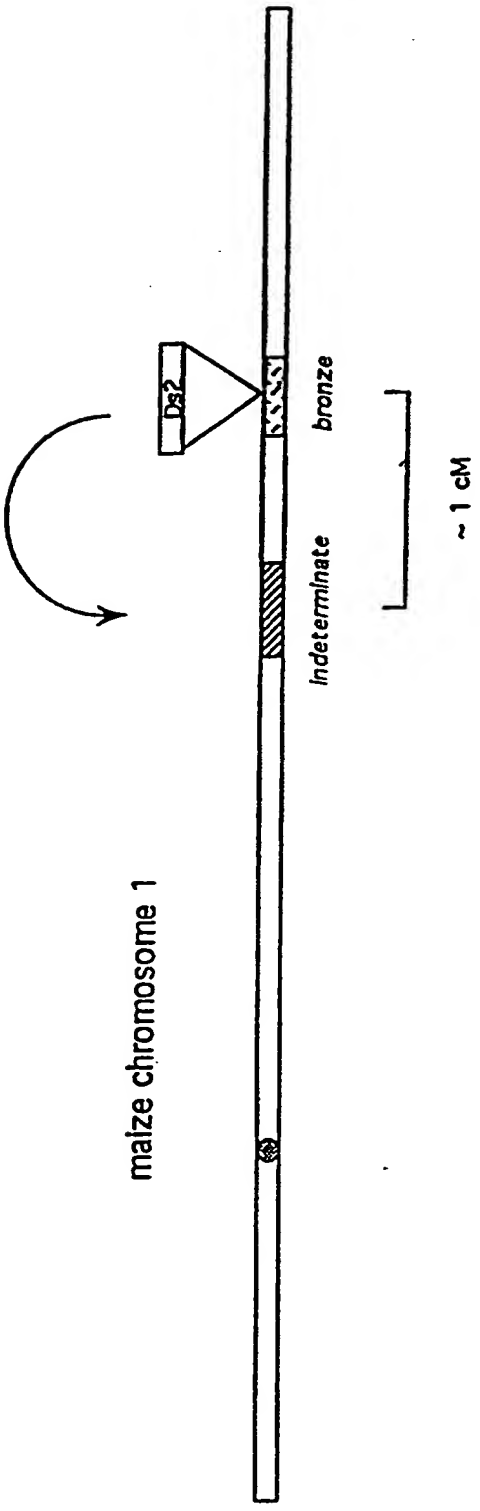


FIGURE 1

1 60

GAGCTCTGGGGGACTTGACTGGGATCAAGAAGCACTTCTCGCGGAAGCACGGGGAGAAGC
GGTGGTGCTGCGAGCGCTGCGGGAAGCCGTACGCCGTGCAGTCGGACTGGAAGGCGCACG
TCAAGGGGTGTGGCACGCGCGAGTACCGCTGCGACTGCGGCATCCTCTTCTCCAGGTACA
TCATCTCATGATCACCGTGCACATATGCATGGACGACGTGTGCTTTGCTGTAATTGTAAA
CGCTGATCATTTTTACTAACAACCATGCTGGATATAATAGCCTAATCTCTCACCGGACGG
ATCGAGAGAAAACCTAGCTAGACGGGATCGATCGGTCCAGCAGGTTGCCGCCGACGACTG
TTCCATCGATCGAGCCTGTTAATTTAGTCATAAAAAGGATCGAGCATATGCATGTATATG
AACTATCTTCCTTCACTGACCAACATCATATCATGCATGGAGCTAGCTAGTTAATCAGTA
CATATACTCCTATATATACATAGGTTTTCAAGAACAGTGGGTGATTCTGAAGCAACCTAA
ATATATATAGATACCAAAAAANATATGAAGTCATCAGCACGATCTGCGAGCGGGTACGGT
TCTTGAACTCTTCTGATGGTTGCAGTAATACCGGCCAACAAAAATATATTATATATTTAT
CGTCCGCTAGTTGATTTTTAAACTAAATGCGCACTGATAAAAAAAGAAGGGTTGGAGTAC
TATATATACAAGAGCATGTGGCCTTCAGTTACAATTTTAGGGTTTCCATGCATCCTGTCA
TAAAACTATTTGCATGATCACATCCCTATATATCGGGATACTACTGTTGTGAAAAAACCA
TGAGTCCCTGGTCAAACCAGTATATGTACATGCAATATGTTTATTGCATGCATATTTGGG
AATGAACATCCTCTGCCTGCACCAACTTTATGGCAGTACGTCCATGTGGCCATCATGACA
CATTCCCTTCAAAAATGGAACATATATAGCTACAGCATATGAAGCAATTGAAGAGTACTT
TAATTGTGAAATAGTACTACTGCAAGTATATATATATGTAGTAGCACAACAGTCGAATAA
TGCAGTGCATTAGATATAGTAGTGAAGTTAAGAGTTAGTTTCCAAATCTTTTACTAGAGA
GAGCATAAAAAATCTATAAAAAATTCTAATTCAACTTCTAATGTATCTTATGTTAAGAAA
GGGGTATATATAAAAAGAGTAAATTCTGTCATTAGATACATCGTTAGCAGTAGTACCACT
GAATTTAATTACGTCCTATACACACGCGCACACACATGCATGCATGCATCTGCATGCTTC
TTTTCAGTAGTGATCACAAAGGAAACTGACAAAAGAACCTAGCTAATCATAGGACGCAGC
TTTTCGTCAGCAAAGTTAAACGAAACTTTACATGCATGGATTGCATTGAGTACTCACGCA
TGTGCACGTCAACACGCGCACACATATAGTATATTAACATAGTACTTTATATACCAACTA
ATTAATAAAGTCATTGACTCCTCTGTCCTCTGGTCATTTGTTTAGCTAATTAACCCGTTT
CGTTTGATGCATGCATGGTCTCTCTGGCGTGGTCGTGCAGGAAGGACAGCCTACTCACGC
ACAGGGCCTTCTGCGATGCCCTAGCAGAGGAGAGCGCGAGGCTTCTTGCAGCAGCAGCTA
ACAACGGCAGCACTATCACCACCACGACCAGCAGCAGCAACAACAATGATCTTCTCAACG
CCAGCAATAATATCACGCCATTATTCCTCCCGTTCGCCAGCTCTCCTCCTCCTGTCGTTG
CGGCGGCACAAAACCCTAATAACACCCTCTTCTTCCTGCGCCAAGATGTCTCCCTTCCTG
CAACCGAGGGTGACAATGCAACAGCAACCCTCGCCCTATCTTGACCTCCATATGCATGTC
GACGCCAGCATCGTCACCACCACCGGTGCTCGCGGACGGCACGCCGGTCAGCTTTGGCCT
CGCTCTGGACGGCTCGGTGGCCACCGTCGCGCACCGGCGCCTCACTAGGGACTTCCTCGG
TGTCGATGGTGGCGGTAGTCAGGTCGAGGAGCTGCAGCTTCCACTGTGCGCCACAGCAGC
AGCAGGTGCCAGCCGCACGCCAGCTGNCCACCGACCTGACAAGGCAGTGCTCGGCGGCCG
GCTGCCGTAACGAGACCTGGAGCCACAACTTCTAGGCCCGCTATATACTTCAAGCTGCAT
TGAGACTTTGAGAGACGAATGAACGGAACACCCGAACTGCATGCACTCTAGCTTGAAGAG
CAACCAAAACTGGAGTAGCAAGTATGGTGCACTACTGTTGTTAATTTACCTTAATTTATT
GATCTCTGGTTAGTTCTGTTTTCATTTAGGGCAATGCGGGCTAGCTAATTAATTCGATGT
GCACAACTTTTGATGAATGGACCATAAAGTTTATCTTGTTGCTTTTTTTTTGTTTGATTA
TGTTTCGCTGCACACCCATGTGTTCTCATAATGGTATGTCGAAAGAAATAGATGATATAC
TAATATAACCATATCAGTCTAAACAACATGAATAAAGATTCAATCAAGAGGAGTGGCACA
TGCATGGTTACTGATGGTGGTACGGAGTCATCGATAAGTGGTAGTGGAGGAAAAGCTTGG
TGCAAACGGCGATGAATACAACGACACGTATAGCACCGTTTAACTTGGATGAAAGACGAC
TCGTCGTGGAAGTTGAGAGCAGTCATGCAAAGAACACTTTCCAAAAACCTTATTAAATAT
GTCCTCTATCTGTGCAAGGTTAGAAAGATGAGAATTATGGAGATCTACTCTCCTGAATCC
TGATTGGTGATGCACGTAAATGCTCAGGATGAAGAGGCTATGACGTCAGTGCAACATTGA
GAAGTGAAAAATACTAATTTATATCTTAAGATTTTTCAAAGTAGGAGCTC

2930

FIGURE 2

```
                                            27                                    54
    GAG CTC TGG GGG ACT TGA CTG GGA TCA AGA AGC ACT TCT CGC GGA AGC ACG GGG
(a) E   L   W   G   T   .   L   G   S   R   S   T   S   R   G   S   T   G
(b)  S   S   G   G   L   D   W   D   Q   E   A   L   L   A   E   A   R   G
(c)   A   L   G   D   L   T   G   I   K   K   H   F   S   R   K   H   G   E

                                            81                                   108
    AGA AGC GGT GGT GCT GCG AGC GCT GCG GGA AGC CGT ACG CCG TGC AGT CGG ACT
    R   S   G   G   A   A   S   A   A   G   S   R   T   P   C   S   R   T
     E   A   V   V   L   R   A   L   R   E   A   V   R   R   A   V   G   L
      K   R   W   C   C   E   R   C   G   K   P   Y   A   V   Q   S   D   W

                                           135                                   162
    GGA AGG CGC ACG TCA AGG GGT GTG GCA CGC GCG AGT ACC GCT GCG ACT GCG GCA
    G   R   R   T   S   R   G   V   A   R   A   S   T   A   A   T   A   A
     E   G   A   R   Q   G   V   W   H   A   R   V   P   L   R   L   R   H
      K   A   H   V   K   G   C   G   T   R   E   Y   R   C   D   C   G   I

                                           189                                   216
    TCC TCT TCT CCA GGT ACA TCA TCT CAT GAT CAC CGT GCA CAT ATG CAT GGA CGA
    S   S   S   P   G   T   S   S   H   D   H   R   A   H   M   H   G   R
     P   L   L   Q   V   H   H   L   M   I   T   V   H   I   C   M   D   D
      L   F   S   R   Y   I   I   S   .   S   P   C   T   Y   A   W   T   T

                                           243                                   270
    CGT GTG CTT TGC TGT AAT TGT AAA CGC TGA TCA TTT TTA CTA ACA ACC ATG CTG
    R   V   L   C   C   N   C   K   R   .   S   F   L   L   T   T   M   L
     V   C   F   A   V   I   V   N   A   D   H   F   Y   .   Q   P   C   W
      C   A   L   L   .   L   .   T   L   I   I   F   T   N   N   H   A   G

                                           297                                   324
    GAT ATA ATA GCC TAA TCT CTC ACC GGA CGG ATC GAG AGA AAA CCT AGC TAG ACG
    D   I   I   A   .   S   L   T   G   R   I   E   R   K   P   S       T
     I   .   .   P   N   L   S   P   D   G   S   R   E   N   L   A   R   R
      Y   N   S   L   I   S   H   R   T   D   R   E   K   T   .   L   D   G
```

FIGURE 3A

```
GGA TCG ATC GGT CCA GCA GGT TGC CGC CGA CGA CTG TTC CAT CGA TCG AGC CTG
 G   S   I   G   P   A   G   C   R   R   R   L   F   H   R   S   S   L
  D   R   S   V   Q   Q   V   A   A   D   D   C   S   I   D   R   A   C
   I   D   R   S   S   R   L   P   P   T   T   V   P   S   I   E   P   V

                                405                                 432
TTA ATT TAG TCA TAA AAA GGA TCG AGC ATA TGC ATG TAT ATG AAC TAT CTT CCT
 L   I   .   S   .   K   G   S   S   I   C   M   Y   M   N   Y   L   P
  .   F   S   H   K   K   D   R   A   Y   A   C   I   .   T   I   F   L
   N   L   V   I   K   R   I   E   H   M   H   V   Y   E   L   S   S   F

                                459                                 486
TCA CTG ACC AAC ATC ATA TCA TGC ATG GAG CTA GCT AGT TAA TCA GTA CAT ATA
 S   L   T   N   I   I   S   C   M   E   L   A   S   .   S   V   H   I
  H   .   P   T   S   Y   H   A   W   S   .   L   V   N   Q   Y   I   Y
   T   D   Q   H   H   I   M   H   G   A   S   .   L   I   S   T   Y   T

                                513                                 540
CTC CTA TAT ATA CAT AGG TTT TCA AGA ACA GTG GGT GAT TCT GAA GCA ACC TAA
 L   L   Y   I   H   R   F   S   R   T   V   G   D   S   E   A   T   .
  S   Y   I   Y   I   G   F   Q   E   Q   W   V   I   L   K   Q   P   K
   P   I   Y   T   .   V   F   K   N   S   G   .   F   .   S   N   L   N

                                567                                 594
ATA TAT ATA GAT ACC AAA AAA NAT ATG AAG TCA TCA GCA CGA TCT GCG AGC GGG
 I   Y   I   D   T   K   K       M   K   S   S   A   R   S   A   S   G
  Y   I   .   I   P   K   X   I   .   S   H   Q   H   D   L   R   A   G
   I   Y   R   Y   Q   K       Y   E   V   I   S   T   I   C   E   R   V

                                621                                 648
TAC GGT TCT TGA ACT CTT CTG ATG GTT GCA GTA ATA CCG GCC AAC AAA AAT ATA
 Y   G   S   .   T   L   L   M   V   A   V   I   P   A   N   K   N   I
  T   V   L   E   L   F   .   W   L   Q   .   Y   R   P   T   K   I   Y
   R   F   L   N   S   S   D   G   C   S   N   T   G   Q   Q   K   Y   I

                                675                                 702
TTA TAT ATT TAT CGT CCG CTA GTT GAT TTT TAA ACT AAA TGC GCA CTG ATA AAA
 L   Y   I   Y   R   P   L   V   D   F   .   T   K   C   A   L   I   K
  Y   I   F   I   V   R   .   L   I   F   K   L   N   A   H   .   .   K
   I   Y   L   S   S   A   S   .   F   L   N   .   M   R   T   D   K   K

                                729                                 756
AAA GAA GGG TTG GAG TAC TAT ATA TAC AAG AGC ATG TGG CCT TCA GTT ACA ATT
 K   E   G   L   E   Y   Y   I   Y   K   S   M   W   P   S   V   T   I
  K   K   G   W   S   T   I   Y   T   R   A   C   G   L   Q   L   Q   F
   R   R   V   G   V   L   Y   I   Q   E   H   V   A   F   S   Y   N   F

                                783                                 810
TTA GGG TTT CCA TGC ATC CTG TCA TAA AAC TAT TTG CAT GAT CAC ATC CCT ATA
 L   G   F   P   C   I   L   S   .   N   Y   L   H   D   H   I   P   I
  .   G   F   H   A   S   C   H   K   T   I   C   M   I   T   S   L   Y
   R   V   S   M   H   P   V   I   K   L   F   A   .   S   H   P   Y   I

                                837                                 864
TAT CGG GAT ACT ACT GTT GTG AAA AAA CCA TGA GTC CCT GGT CAA ACC AGT ATA
 Y   R   D   T   T   V   V   K   K   P   .   V   P   G   Q   T   S   I
  I   G   I   L   L   L   .   K   N   H   E   S   L   V   K   P   V   Y
   S   G   Y   Y   C   C   E   K   T   M   S   P   W   S   N   Q   Y   M
```

FIGURE 3B

891 918
TGT ACA TGC AAT ATG TTT ATT GCA TGC ATA TTT GGG AAT GAA CAT CCT CTG CCT
C T C N M F I A C I F G N E H P L P
V H A I C L L H A Y L G M N I L C L
Y M Q Y V Y C M H I W E . T S S A C

945 972
GCA CCA ACT TTA TGG CAG TAC GTC CAT GTG GCC ATC ATG ACA CAT TCC CTT CAA
A P T L W Q Y V H V A I M T H S L Q
H Q L Y G S T S M W P S . H I S P F K
T N F M A V R P C G H H D T F P S K

999 1026
AAA TGG AAC ATA TAT AGC TAC AGC ATA TGA AGC AAT TGA AGA GTA CTT TAA TTG
K W N I Y S Y S I . S N . R V L . L
N G T Y I A T A Y E A I E E Y F N C
M E H I . L Q H M K Q L K S T L I V

1053 1080
TGA AAT AGT ACT ACT GCA AGT ATA TAT ATA TGT AGT AGC ACA ACA GTC GAA TAA
. N S T T A S I Y I C S S T T V E .
E I V L L Q V Y I Y V V A Q Q S N N
K . Y Y C K Y I Y M . . H N S R I M

1107 1134
TGC AGT GCA TTA GAT ATA GTA GTG AAG TTA AGA GTT AGT TTC CAA ATC TTT TAC
C S A L D I V V K L R V S F Q I F Y
A V H . I . . . S . E L V S K S F T
Q C I R Y S S E V K S . F P N L L L

1161 1188
TAG AGA GAG CAT AAA AAA TCT ATA AAA AAT TCT AAT TCA ACT TCT AAT GTA TCT
. R E H K K S I K N S N S T S N V S
R E S I K N L . K I L I Q L L M Y L
E R A . K I Y K K F . F N F . C I L

1215 1242
TAT GTT AAG AAA GGG GTA TAT ATA AAA AGA GTA AAT TCT GTC ATT AGA TAC ATC
Y V K K G V Y I K R V N S V I R Y I
M L R K G Y I . K E . I L S L D T S
C . E R G I Y K K S K F C H . I H R

1269 1296
GTT AGC AGT AGT ACC ACT GAA TTT AAT TAC GTC CTA TAC ACA CGC GCA CAC ACA
V S S S T T E F N Y V L Y T R A H T
L A V V P L N L I T S Y T H A H T
. Q . Y H . I . L R P I H T R T H M

1323 1350
TGC ATG CAT GCA TCT GCA TGC TTC TTT TCA GTA GTG ATC ACA AAG GAA ACT GAC
C M H A S A C F F S V V I T K E T D
A C M H L H A S F Q . . S Q R K L T
H A C I C M L L F S S D H K G N . Q

1377 1404
AAA AGA ACC TAG CTA ATC ATA GGA CGC AGC TTT TCG TCA GCA AAG TTA AAC GAA
K R T . L I I G R S F S S A K L N E
K E P S . S . D A A F R Q Q S . T K
K N L A N H R T Q L F V S K V K R N

FIGURE 3C

```
                                         1431                                      1458
ACT TTA CAT GCA TGG ATT GCA TTG AGT ACT CAC GCA TGT GCA CGT CAA CAC GCG
 T   L   H   A   W   I   A   L   S   T   H   A   C   A   R   Q   H   A
  L   Y   M   H   G   L   H   .   V   L   T   H   V   H   V   N   T   R
   F   T   C   M   D   C   I   E   Y   S   R   M   C   T   S   T   R   A

                                         1485                                      1512
CAC ACA TAT AGT ATA TTA ACA TAG TAC TTT ATA TAC CAA CTA ATT AAT AAA GTC
 H   T   Y   S   I   L   T   .   Y   F   I   Y   Q   L   I   N   K   V
  T   H   I   V   Y   .   H   S   T   L   Y   T   N   .   L   I   K   S
   H   I   .   Y   I   N   I   V   L   Y   I   P   T   N   .   S   .   S   H

                                         1539                                      1566
ATT GAC TCC TCT GTC CTC TGG TCA TTT GTT TAG CTA ATT AAC CCG TTT CGT TTG
 I   D   S   S   V   L   W   S   F   V   .   L   I   N   P   F   R   L
  L   T   P   L   S   S   G   H   L   F   S   .   L   T   R   F   V   .
   .   L   L   C   P   L   V   I   C   L   A   N   .   P   V   S   F   D

                                         1593                                      1620
ATG CAT GCA TGG TCT CTC TGG CGT GGT CGT GCA GGA AGG ACA GCC TAC TCA CGC
 M   H   A   W   S   L   W   R   G   R   A   G   R   T   A   Y   S   R
  C   M   H   G   L   S   G   V   V   V   Q   E   G   Q   P   T   H   A
   A   C   M   V   S   L   A   W   S   C   R   K   D   S   L   L   T   H

                                         1647                                      1674
ACA GGG CCT TCT GCG ATG CCC TAG CAG AGG AGA GCG CGA GGC TTC TTG CAG CAG
 T   G   P   S   A   M   P   .   Q   R   R   A   R   G   F   L   Q   Q
  Q   G   L   L   R   C   P   S   R   G   E   R   E   A   S   C   S   S
   R   A   F   C   D   A   L   A   E   E   S   A   R   L   L   A   A   A

                                         1701                                      1728
CAG CTA ACA ACG GCA GCA CTA TCA CCA CCA CGA CCA GCA GCA GCA ACA ACA ATG
 Q   L   T   T   A   A   L   S   P   P   R   P   A   A   A   T   T   M
  S   .   Q   R   Q   H   Y   H   H   H   D   Q   Q   Q   Q   Q   Q   .
   A   N   N   G   S   T   I   T   T   T   T   S   S   S   N   N   N   D

                                         1755                                      1782
ATC TTC TCA ACG CCA GCA ATA ATA TCA CGC CAT TAT TCC TCC CGT TCG CCA GCT
 I   F   S   T   P   A   I   I   S   R   H   Y   S   S   R   S   P   A
  S   S   Q   R   Q   Q   .   Y   H   A   I   I   P   P   V   R   Q   L
   L   L   N   A   S   N   N   I   T   P   L   F   L   P   F   A   S   S

                                         1809                                      1836
CTC CTC CTC CTG TCG TTG CGG CGG CAC AAA ACC CTA ATA ACA CCC TCT TCT TCC
 L   L   L   L   S   L   R   R   H   K   T   L   I   T   P   S   S   S
  S   S   S   C   R   C   G   G   T   K   P   .   .   H   P   L   L   P
   P   P   P   V   V   A   A   A   Q   N   P   N   N   T   L   F   F   L

                                         1863                                      1890
TGC GCC AAG ATG TCT CCC TTC CTG CAA CCG AGG GTG ACA ATG CAA CAG CAA CCC
 C   A   K   M   S   P   F   L   Q   P   R   V   T   M   Q   Q   Q   P
  A   P   R   C   L   P   S   C   N   R   G   .   Q   C   N   S   N   P
   R   Q   D   V   S   L   P   A   T   E   G   D   N   A   T   A   T   L

                                         1917                                      1944
TCG CCC TAT CTT GAC CTC CAT ATG CAT GTC GAC GCC AGC ATC GTC ACC ACC ACC
 S   P   Y   L   D   L   H   M   H   V   D   A   S   I   V   T   T   T
  R   P   I   L   T   S   I   C   M   S   T   P   A   S   S   P   P   P
   A   L   S   .   P   P   Y   A   C   R   E   Q   H   R   H   H   H   R
```

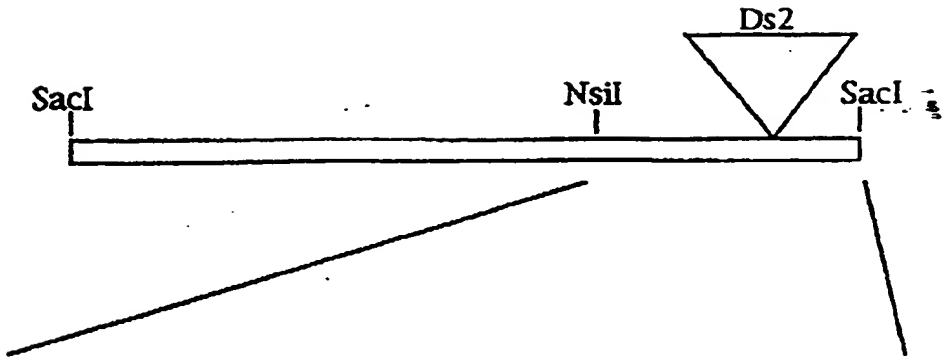
FIGURE 3D

```
                                  1971                                 1998
GGT GCT CGC GGA CGG CAC GCC GGT CAG CTT TGG CCT CGC TCT GGA CGG CTC GGT
G   A   R   G   R   H   A   G   Q   L   W   P   R   S   G   R   L   G
 V   L   A   D   G   T   P   V   S   F   G   L   A   L   D   G   S   V
  C   S   R   T   A   R   R   S   A   L   A   S   L   W   T   A   R   W

                                  2025                                 2052
GGC CAC CGT CGC GCA CCG GCG CCT CAC TAG GGA CTT CCT CGG TGT CGA TGG TGG
G   H   R   R   A   P   A   P   H   .   G   L   P   R   C   R   W   W
 A   T   V   A   H   R   R   L   T   R   D   F   L   G   V   D   G   G
  P   P   S   R   T   G   A   S   L   G   T   S   S   V   S   M   V   A

                                  2079                                 2106
CGG TAG TCA GGT CGA GGA GCT GCA GCT TCC ACT GTG CGC CAC AGC AGC AGC AGG
R   .   S   G   R   G   A   A   A   S   T   V   R   H   S   S   S   R
 G   S   Q   V   E   E   L   Q   L   P   L   C   A   T   A   A   A   G
  V   V   R   S   R   S   C   S   F   H   C   A   P   Q   Q   Q   Q   V

                                  2133                                 2160
TGC CAG CCG CAC GCC AGC TGN CCA CCG ACC TGA CAA GGC AGT GCT CGG CGG CCG
C   Q   P   H   A   S   X   P   P   T   .   Q   G   S   A   R   R   P
 A   S   R   T   P   A       H   R   P   D   K   A   V   L   G   G   R
  P   A   A   R   Q   L       T   D   L   T   R   Q   C   S   A   A   G

                                  2187                                 2214
GCT GCC GTA ACG AGA CCT GGA GCC ACA ACT TCT AGG CCC GCT ATA TAC TTC AAG
A   A   V   T   R   P   G   A   T   T   S   R   P   A   I   Y   F   K
 L   P   .   R   D   L   E   P   Q   L   L   G   P   L   Y   T   S   S
  C   R   N   E   T   W   S   H   N   F   .   A   R   Y   I   L   Q   A

                                  2241                                 2268
CTG CAT TGA GAC TTT GAG AGA CGA ATG AAC GGA ACA CCC GAA CTG CAT GCA CTC
L   H   .   D   F   E   R   R   M   N   G   T   P   E   L   H   A   L
 C   I   E   T   L   R   D   E   .   T   E   H   P   N   C   M   H   S
  A   L   R   L   .   E   T   N   E   R   N   T   R   T   A   C   T   L

                                  2295                                 2322
TAG CTT GAA GAG CAA CCA AAA CTG GAG TAG CAA GTA TGG TGC ACT ACT GTT GTT
.   L   E   E   Q   P   K   L   E   .   Q   V   W   C   T   T   V   V
 S   L   K   S   N   Q   N   W   S   S   K   Y   G   A   L   L   L   L
  A   .   R   A   T   K   T   G   V   A   S   M   V   H   Y   C   C   .

                                  2349                                 2376
AAT TTA CCT TAA TTT ATT GAT CTC TGG TTA GTT CTG TTT TCA TTT AGG GCA ATG
N   L   P   .   F   I   D   L   W   L   V   L   F   S   F   R   A   M
 I   Y   L   N   L   L   I   S   G   .   F   C   F   H   L   G   Q   C
  F   T   L   I   Y   .   S   L   V   S   S   V   F   I   .   G   N   A

                                  2403                                 2430
CGG GCT AGC TAA TTA ATT CGA TGT GCA CAA CTT TTG ATG AAT GGA CCA TAA AGT
R   A   S   .   L   I   R   C   A   Q   L   L   M   N   G   P   .   S
 G   L   A   N   .   F   D   V   H   N   F   .   .   M   D   H   K   V
  G   .   L   I   N   S   M   C   T   T   F   D   E   W   T   I   K   F

                                  2457                                 2484
TTA TCT TGT TGC TTT TTT TTT GTT TGA TTA TGT TTC GCT GCA CAC CCA TGT GTT
L   S   C   C   F   F   F   V   .   L   C   F   A   A   H   P   C   V
 Y   L   V   A   F   F   L   F   D   Y   V   S   L   H   T   H   V   F
  I   L   L   L   F   F   C   L   I   M   F   R   C   T   P   M   C   S
```

FIGURE 3E

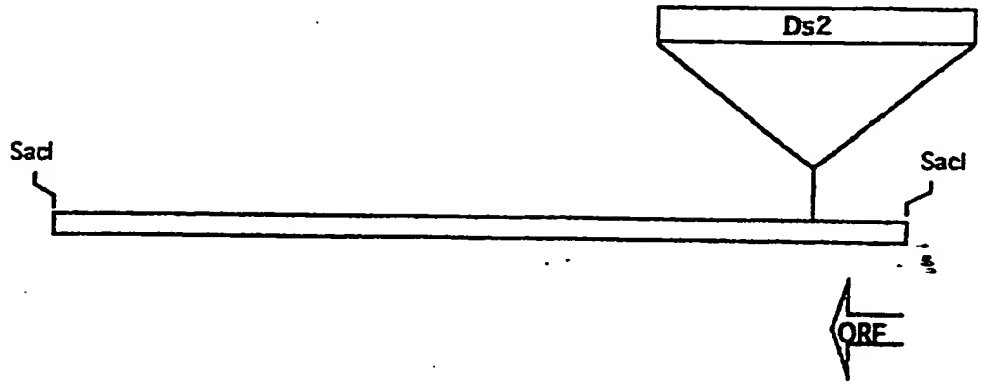
```
                                                2511                                                2538
CTC ATA ATG GTA TGT CGA AAG AAA TAG ATG ATA TAC TAA TAT AAC CAT ATC AGT
L   I   M   V   C   R   K   K   .   M   I   Y   .   Y   N   H   I   S
 S   .   W   Y   V   E   R   N   R   .   Y   T   N   I   T   I   S   V
  H   N   G   M   S   K   E   I   D   D   I   L   I   .   P   Y   Q   S

                                                2565                                                2592
CTA AAC AAC ATG AAT AAA GAT TCA ATC AAG AGG AGT GGC ACA TGC ATG GTT ACT
L   N   N   M   N   K   D   S   I   K   R   S   G   T   C   M   V   T
 .   T   T   .   I   K   I   Q   S   R   G   V   A   H   A   W   L   L
  K   Q   H   E   .   R   F   N   Q   E   E   W   H   M   H   G   Y   .

                                                2619                                                2646
GAT GGT GGT ACG GAG TCA TCG ATA AGT GGT AGT GGA GGA AAA GCT TGG TGC AAA
D   G   G   T   E   S   S   I   S   G   S   G   G   K   A   W   C   K
 M   V   V   R   S   H   R   .   V   V   V   E   E   K   L   G   A   N
  W   W   Y   G   V   I   D   K   W   .   W   R   K   S   L   V   Q   T

                                                2673                                                2700
CGG CGA TGA ATA CAA CGA CAC GTA TAG CAC CGT TTA ACT TGG ATG AAA GAC GAC
R   R   .   I   Q   R   H   V   .   H   R   L   T   W   M   K   D   D
 G   D   E   Y   N   D   T   Y   S   T   V   .   L   G   .   K   T   T
  A   M   N   T   T   T   R   I   A   P   F   N   L   D   E   R   R   L

                                                2727                                                2754
TCG TCG TGG AAG TTG AGA GCA GTC ATG CAA AGA ACA CTT TCC AAA AAC CTT ATT
S   S   W   K   L   R   A   V   M   Q   R   T   L   S   K   N   L   I
 R   R   G   S   .   E   Q   S   C   K   E   H   F   P   K   T   L   L
  V   V   E   V   E   S   S   H   A   K   N   T   F   Q   K   P   Y   .

                                                2781                                                2808
AAA TAT GTC CTC TAT CTG TGC AAG GTT AGA AAG ATG AGA ATT ATG GAG ATC TAC
K   Y   V   L   Y   L   C   K   V   R   K   M   R   I   M   E   I   Y
 N   M   S   S   I   C   A   R   L   E   R   .   E   L   W   R   S   T
  I   C   P   L   S   V   Q   G   .   K   D   E   N   Y   G   D   L   L

                                                2835                                                2862
TCT CCT GAA TCC TGA TTG GTG ATG CAC GTA AAT GCT CAG GAT GAA GAG GCT ATG
S   P   E   S   .   L   V   M   H   V   N   A   Q   D   E   E   A   M
 L   L   N   P   D   W   .   C   T   .   M   L   R   M   K   R   L   .
  S   .   I   L   I   G   D   A   R   K   C   S   G   .   R   G   Y   D

                                                2889                                                2916
ACG TCA GTG CAA CAT TGA GAA GTG AAA AAT ACT AAT TTA TAT CTT AAG ATT TTT
T   S   V   Q   H   .   E   V   K   N   T   N   L   Y   L   K   I   F
 R   Q   C   N   I   E   K   .   K   I   L   I   Y   I   L   R   F   F
  V   S   A   T   L   R   S   E   K   Y   .   F   I   S   .   D   F   S

CAA AGT AGG AGC TC
Q   S   R   S
 K   V   G   A
  K   .   E   L
```

FIGURE 3F




```
ATGCATATGC TCGATCCTTT TTATGACTAA ATTAACAGGC TCGATCGATG      50
GAACAGTCGT CGGCGGCAAC CTGCTGGACC GATCGATCCC GTCTAGCTAG     100
GTTTTCTCTC GATCCGTCCG GTGAGAGATT AGGCTATTAT ATCCAGCATG     150
GTTGTTAGTA AAAATGATCA GCGTTTACAA TTACAGCAAA GCACACGTCG     200
TCCATGCATA TGTGCACGGT GATCATGAGA TGATGTACCT GGAGAAGAGG     250
ATGCCGCAGT CGCAGCGTAC TCGCGCGTGC CACACCCCTT GACGTGCGCC     300
TTCCAGTCCG ACTGCACGGC GTACGCTTCC CGCAGCGCTC GCAGCACCAC     350
CGCTTCTCCC CGTGCTTCCG CGAGAAGTGC TTCTTGATCC CAGTCAAGTC     400
CCCCAGAGCT C                                               411
```

FIGURE 4



ALGDLTGIKK HFSRKHGEKR WCCERCGKPY AVQSDWKAHV KGCGTREYRC 50
DCGILFSRYI IS 62

ORF (Open Reading Frame): encodes protein homologous to zinc-finger regulatory protein

FIGURE 5

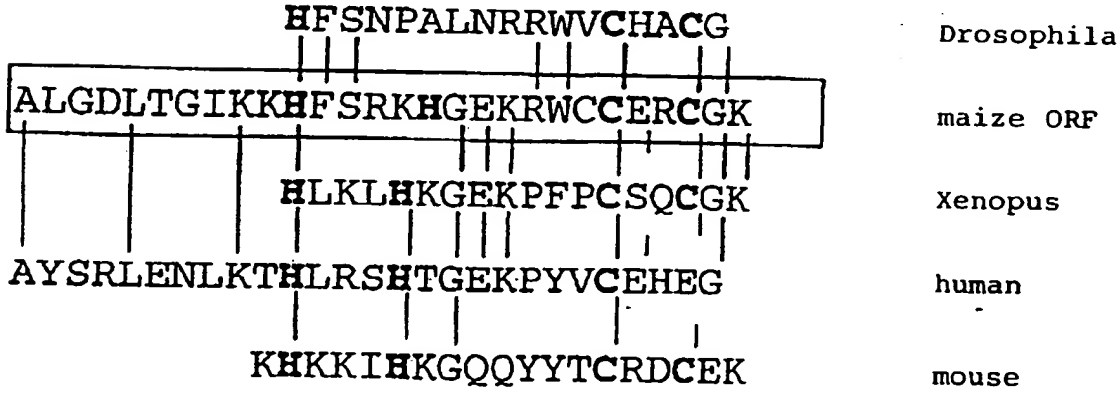


FIGURE 6

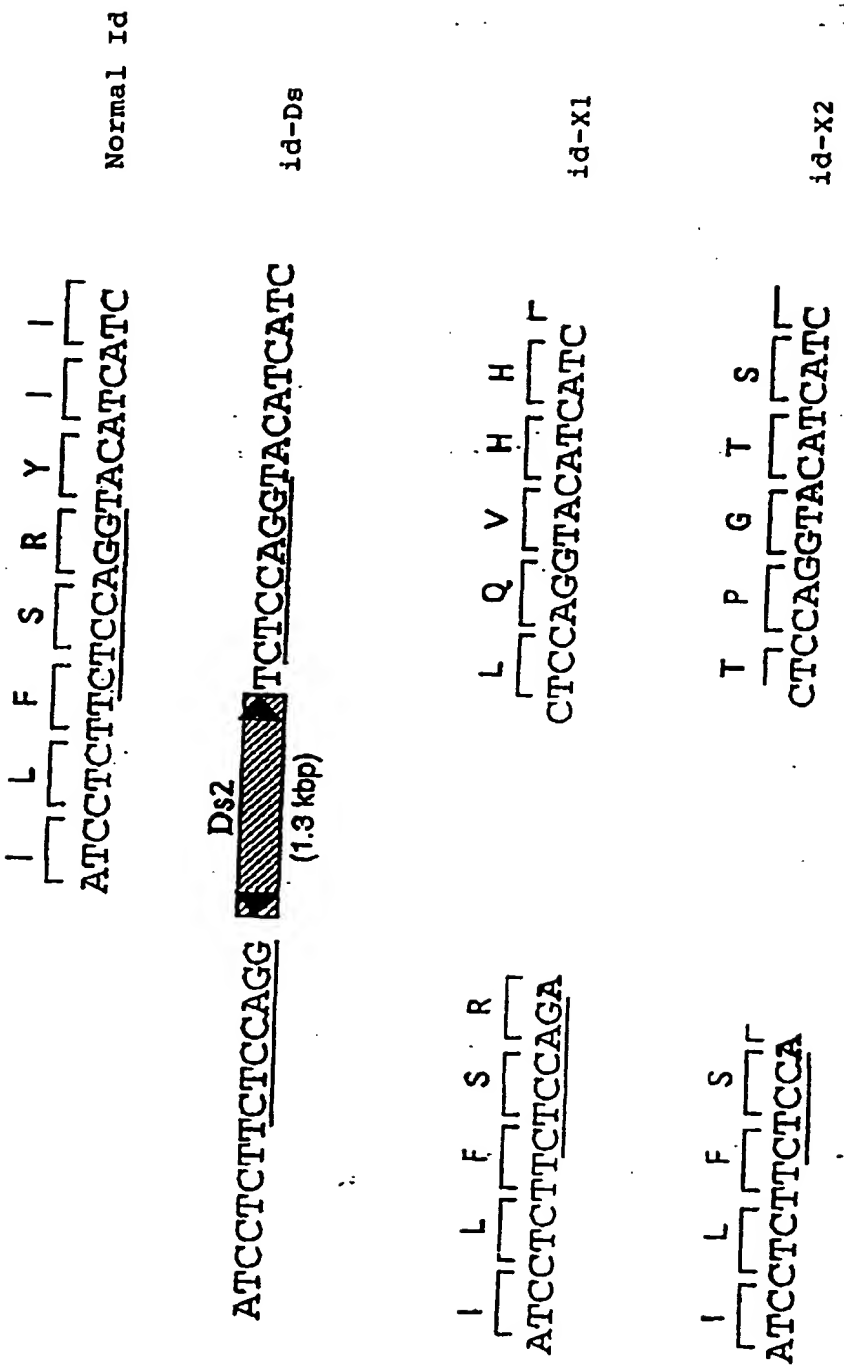

I L F S R Y I I
ATCCTCTTCTCCAGGTACATCATC
Normal Id
Ds2
(1.3 kbp)
ATCCTCTTCTCCAGG
TCTCCAGGTACATCATC
id-Ds
I L F S R
ATCCTCTTCTCCAGA
L Q V H H
CTCCAGGTACATCATC
id-X1
I L F S
ATCCTCTTCTCCA
T P G T S
CTCCAGGTACATCATC
id-X2

FIGURE 7

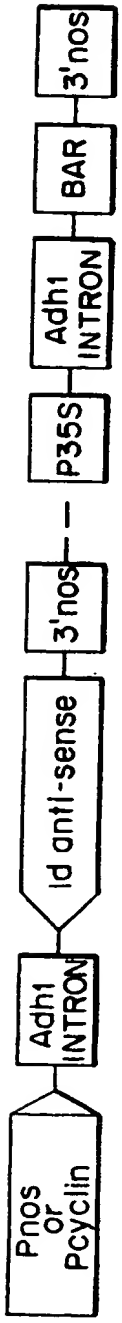


FIG. 8A

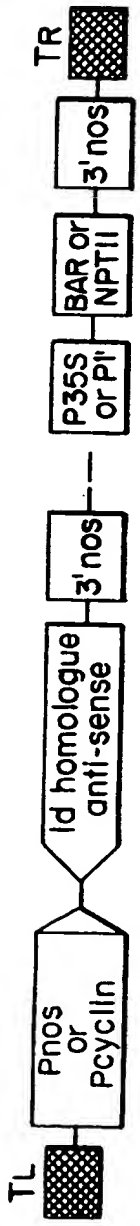


FIG. 8B

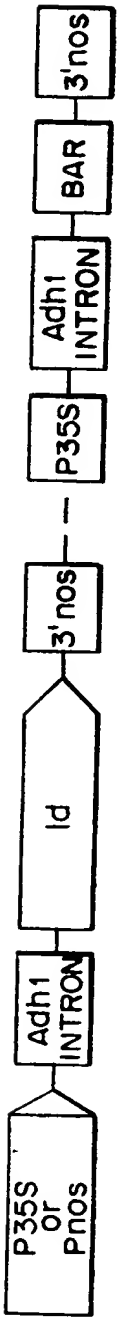


FIG. 9A

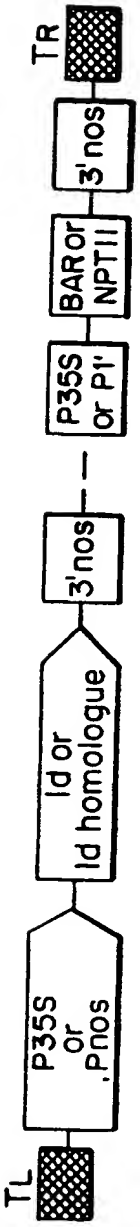


FIG. 9B

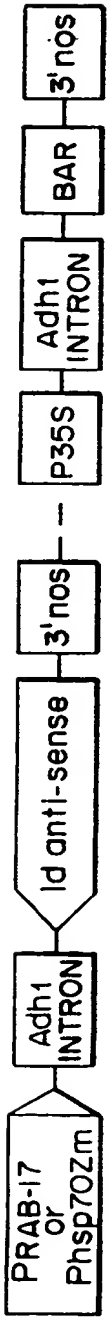


FIG. 10A

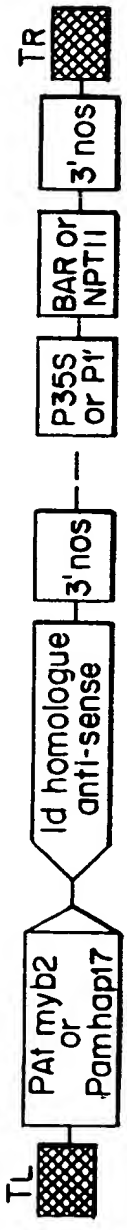


FIG. 10B

FIG. IIA

* GB = GAL4 Binding Site (17mers as described in Ma, J. et al., supra)

** M = Minimal Promoter (TATA Box)

FIG. IIB

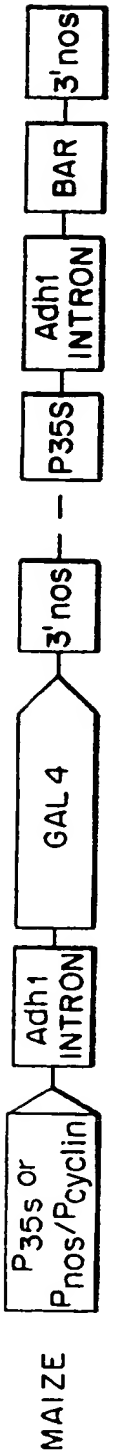


FIG. IIC

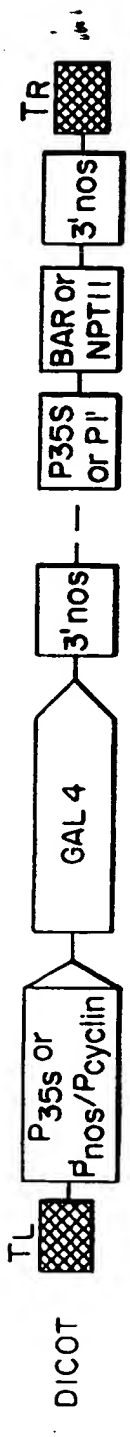


FIG. IID